

Running head: STUDENTS' AGENTIC AND EFFICACIOUS BEHAVIORS AS AN
INDICATOR OF TEACHER EFFECTIVENESS

STUDENTS' AGENTIC AND EFFICACIOUS BEHAVIORS IN THE ELEMENTARY
CLASSROOM AS AN INDICATOR OF TEACHER EFFECTIVENESS IN HIGH POVERTY
SCHOOLS: A CONCURRENT TRIANGULATION MIXED METHODS STUDY.

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF EDUCATION
BY
SARAH WEIMER
DISSERTATION ADVISOR: DR. SERENA SALLOUM

BALL STATE UNIVERSITY
MUNCIE, INDIANA
MAY, 2019

STUDENTS' AGENTIC AND EFFICACIOUS BEHAVIORS IN THE ELEMENTARY
CLASSROOM AS AN INDICATOR OF TEACHER EFFECTIVENESS IN HIGH POVERTY
SCHOOLS: A CONCURRENT TRIANGULATION MIXED METHODS STUDY.

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF EDUCATION
BY
SARAH WEIMER

DISSERTATION ADVISOR: DR. SERENA SALLOUM

APPROVED BY:

Committee Chairperson

Date

Committee Member

Date

Committee Member

Date

Committee Member

Date

Dean of Graduate School

Date

BALL STATE UNIVERSITY
MUNCIE, INDIANA
MAY, 2019

ABSTRACT

DISSERTATION: Students' agentic and efficacious behaviors in the elementary classroom as an indicator of teacher effectiveness in high poverty schools; a concurrent triangulation mixed methods study.

STUDENT: Sarah Weimer

DEGREE: Doctorate of Education

COLLEGE: Teachers College

DATE: May, 2019

PAGES: 249

This concurrent multi-phase mixed methods study investigated whether indicators of Social Cognitive Theory could be observed in Grades 3-5 classrooms and if these indicators could be used to describe teacher effectiveness in high poverty schools. The conceptual framework drew upon Bandura's (1997) measurement of the individual self-efficacy of students and the collective teacher efficacy of the organization to study the level of agency and self-efficacy cultivated by the teacher and students in the classroom environment. Major findings confirmed a correlation between self-efficacy and student outcomes in ELA and mathematics. Looping, the practice of students being paired with the same teacher for multiple years, was also a significant factor. Agentic and self-efficacious student behaviors were observable in classrooms. An observation tool developed over the course of the study was found to be mostly reliable and had internal consistency with the exception of one subscale in mathematics. Preliminary evidence suggests there are observable cultures of student agency and self-efficacy in classrooms, but the sample group and scope of the study were limited. Further research in this area is warranted.

DEDICATION

I dedicate this work to my family.

To my husband, Daniel, thank you for your never-ending support and love, and for always believing in me. You are a constant source of inspiration and a role model for my pursuit of excellence. I could not ask for a better partner and love in this crazy life, and I know I could not have done this without you.

To my children, Olivia, Aidan, and Brennan. Thank you for your patience and understanding when I could not play or was only half listening because I was “working on my dissertation.” You are my inspiration, and I hope I make you a fraction as proud as you make me.

To my parents, Bernie and Kathy, thank you for rearing me to be an independent and empathetic woman, and for never saying “no” when I wanted to buy a new book or learn something new. I (literally) wouldn’t be here without you.

Finally, I dedicate this work to my professional family - the students and educators I’ve had the privilege of working with over the course of my career. Each and every one of you has taught me something about myself, teaching, and learning, and I am forever grateful for those lessons. I especially dedicate this to those students who most needed someone to believe in them and to help them believe in themselves.

ACKNOWLEDGEMENTS

I am incredibly indebted to my dissertation chair, Dr. Serena Salloum, for her guidance, feedback, patience, and encouragement. I am also grateful to Dr. Marilynn Quick who served as my periodic surrogate chair and provided an abundance of support and motivation during this process. Thank you both for helping me succeed in this endeavor.

Thank you to Dr. Kianre Eouanzoui for your guidance on my quantitative analysis and your patience with all of my questions. Thank you to Drs. Kendra Lowery and Lisa Rubenstein for giving your time by agreeing to sit on my committee, and for your wonderful feedback and advice on my proposal and final dissertation.

Finally, I owe a debt of gratitude to Marbelys, Marian, Michelle, and Sarah for their help in the field along the way. I sincerely thank you all, and hope you are able to recognize your enormous contributions to this work.

TABLE OF CONTENTS

ABSTRACT	3
DEDICATION	4
ACKNOWLEDGEMENTS	5
LIST OF FIGURES	12
INTRODUCTION	14
The Problem.....	20
Purpose of the Study	20
Research Questions.....	22
Significance	23
Delimitations.....	24
Definition of Terms	25
Organization of the Study	27
LITERATURE REVIEW	29
Conceptual Framework Preview – The Culture of Agency and Efficacy	30
A Review of Teacher Evaluation Literature	35
Process for Improvement	36
Measurement Tools	38
Practical Measures	40
Implications for This Study.....	44
Social Cognitive Theory – Agency and Efficacy	45
Sources of Self-Efficacy	47
Further Differentiating Efficacy.....	51

Correlations Between Academic Efficacy and Student Achievement	60
Correlations with the Learning Environment.....	64
Correlations with Peer Relationships	65
Implications for this Study	66
RESEARCH METHODS	69
Research Questions	70
Research Design	70
Description of the Sample.....	74
Demographics	76
Instruments.....	76
CORE District SEL/MESH Survey	77
PALS Student Survey	79
CASE Observation Tool	81
Data Collection	83
Phase I – Quantitative Collection.....	84
Phase II – Qualitative Collection	86
Data Analysis	88
Phase I - Quantitative Analysis	88
Phase II - Qualitative Analysis.....	90
Limitations	91
Summary	92
RESULTS	94
Research Questions	95

Description of the Sample.....	95
Participant Demographics	97
Results.....	99
Demographics and Achievement	100
Demographics and Survey Results.....	104
Phase I Quantitative Research Question 1	106
Phase II Research Question 2	109
Phase II Research Question 3	161
Overall Question 4	166
Summary	166
DISCUSSION.....	167
Purpose of the Study	167
Research Questions.....	168
Review of the Research Methods	168
Findings Related to the Literature	169
Demographics	169
Quantitative Phase: Research Question One.....	171
Qualitative Phase: Research Question Two	172
Qualitative Phase: Research Question Three	179
Research Question Four	183
Limitations	185
Implications and Recommendations for Practice	186
Recommendations for Teachers	187

Recommendations for Leaders.....	189
Implications for Theory	189
Recommendations for Further Research.....	190
REFERENCES	192
Appendix A	219
Appendix B.....	221
Appendix C.....	223
Appendix D	225
Appendix E	227
Appendix F	229
Appendix G	230
Appendix H	234
Appendix I	240
Appendix J.....	243
Appendix K	246

LIST OF TABLES

Table 1. Examples of the Levels of Specificity and Their Correlation to Achievement.....	61
Table 2. Demographic Population Numbers [and Percent] of Garfield Academy and Washington Academy Enrolled in 3rd, 4th, and 5th grades	76
Table 3. Reliability of CORE District MESH Survey Competencies Using Cronbach's Alpha ..	78
Table 4. Reliability of PALS Teacher Mastery Goal and Academic Press Survey Subscales Using Cronbach's Alpha	80
Table 5. Demographic Population Numbers [and Percent] of Garfield Academy and Washington Academy in 3rd, 4th, and 5th grades, and the Sample Groups	98
Table 6. Demographic Percentages of Participants by Classroom at Garfield and Washington Academies	99
Table 7. Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA and Mathematics by Looped Status	102
Table 8. Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA for Students According to Special Education Enrollment	103
Table 9. Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA and Mathematics Based on School Enrollment.....	104
Table 10. Results of a Correlation Analysis of the Students' Acuity C ELA and Mathematics Outcomes with the Significant Categorical Survey Indicators from the February/March Collection	107
Table 11. Results of a Correlation Analysis of the Categorical Survey Indicators from the December Collection	107

Table 12. Results of a Correlation Analysis of the Categorical Survey Indicators from the February/March Collection	108
Table 13. Predictors of Acuity C Outcomes in ELA.....	108
Table 14. Predictors of Acuity C Outcomes in Mathematics	109
Table 15. Open Codes with Example Data from One Round of 15-Minute Observations in the 10 Participating Classrooms	111
Table 16. Results of a Correlation Analysis of the Observation Category Subscales in ELA	162
Table 17. Reliability of CASE Observation Indicators Overall and as Subscale Categories in ELA, Mathematics and Combined as Measured by Cronbach's Alpha	163
Table 18. Means, Standard Deviations, and Confidence Intervals of Acuity A outcomes in ELA and Mathematics by Demographic Variable	234
Table 19. Means, Standard Deviations, and Confidence Intervals of Acuity B Outcomes in ELA and Mathematics by Demographic Variable	235
Table 20. Means, Standard Deviations, and Confidence Intervals of Acuity C Outcomes in ELA and Mathematics by Demographic Variable	236
Table 21. Means, Standard Deviations, and Confidence Intervals of December Categorical Survey Results by Demographic Variable.	237
Table 22. Means, Standard Deviations, and Confidence Intervals of February/March Categorical Survey Results by Demographic Variable.	238
Table 23. Means, Standard Deviations, and Confidence Intervals of Teacher Perception of Student (TPS) Survey Results by Demographic Variable.....	239

LIST OF FIGURES

Figure 1. The relationship of the Culture of Agency and Self-Efficacy (CASE) conceptual framework to self and collective efficacy.	31
Figure 2. The CORE District SEL survey results for grades 4-12 as compared to academic and behavioral outcomes as it appears in Krachman, Arnold, and Larocca (2016, p. 15).....	54
Figure 3. “The phases and subprocesses of self-regulation” as it appears in Zimmerman (2002, p. 67) from J. E. Davison and R. Sternberg (Eds) (2003).	56
Figure 4. The concurrent triangulation design with an instrument-building, sequential exploratory model embedded in the qualitative phase.....	71
Figure 5. The timing, process and structure of Phase I (quantitative) and Phase II (qualitative) data collection.....	84
Figure 6. The result of the research process on the final iteration of the observation instrument, Tool 3, as presented in Appendix F.....	114
Figure 7. Examples of symmetrical patterns created by groups of students in Classroom 3..	123
Figure 8. Students in Classroom 7 turned their mathematics work into leveled folders posted on a side bulletin board based on their perceived level of progress towards their success on the task.....	125
Figure 9. Students in several classrooms turned their work into leveled bins students, such as this one found in Classroom 2.	126
Figure 10. Student work hung on the bulletin board showing students’ self assessment and justification for the score.	127
Figure 11. Students in Classroom 5 write equivalent fractions for decimal numbers on their desks using dry erase markers.	130

Figure 12. Students tracked their achievement on the Acuity assessments relative to their goals and the class average in a bar graph.	134
Figure 13. Evidence from Classroom 3 of an exemplar quiz displayed on the front board identifying the computational error in yellow, and checking the correct use of methods and computation for the remaining problems.	138
Figure 14. Students in some classrooms were provided with rubrics and conversation starters to help provided targeted feedback and critique to their peers.	142
Figure 15. A sample student progress report from Classroom 4.	143
Figure 16. The learning goals for comprehension, accuracy, fluency, and expand vocabulary were listed on posters on the wall in the front of Classroom 6.	145
Figure 17. A student's draft of his goal sheet for his portfolio along with his drafts and final copy of his corresponding writing piece containing errors with commas.	147
Figure 18. Students in Classroom 4 used these growth plan sheets using the highlight goals from their progress reports to set goals for the next quarter.	148
Figure 19. CHOW tracking sticker charts from Classroom 2.	149

CHAPTER 1

INTRODUCTION

As a part of the Civil Rights Act of 1964, the seminal Coleman Report first examined the relationship between poverty and teaching, learning, and the overall academic achievement of students. President Lyndon B. Johnson commissioned the National Center for Educational Statistics to examine the effects segregation and inequalities in education had on outcomes for students (Coleman, 1966). Ultimately Coleman and his team found achievement inequities, particularly with African Americans. They explained that poverty and the educational attainment levels of parents likely caused the disparities in achievement. Coleman noted these factors put poor children at nearly unrecoverable disadvantages in both verbal and nonverbal skills. These gaps continued to grow, and when these students were in twelfth grade, they were nearly three and a half years behind their more affluent peers (Coleman, 1966).

Coleman's findings continue to resonate in contemporary education, as minority students who grow up in generational high-poverty and live in urban environments are some of the hardest to teach and reach. This is due in large part to historical discriminatory practices and issues affecting academics that accompany poverty, such as stress, trauma, instability, hunger, violence, poor health, mental illness, addiction, trafficking, etc. (Berliner, 2013; Socol, 2014). Because disadvantaged students generally come from families who also grew up in poverty, students in generational high-poverty often lack the familial support systems that provide the knowledge, skills, and resources to cope, overcome, and thrive. These students therefore need social services to help break the cycle of poverty and all that it entails, and schools with high levels of students with low socio-economic status (SES) must position themselves to be a very substantial and powerful part of that system of services.

Facilitating Achievement and “Soft Skills” in Low SES Schools

Importance of Developing Student Agency and Efficacy. The ability for children in poverty to be academically competitive with their affluent peers has become increasingly important, as the US economy has evolved over the last 20 years (Farrington et al., 2012). Fewer low skill jobs are available in the United States as the global economy expands, and more domestic employers are demanding workers who have attained greater levels of education and competence in “soft skills” such as collaboration, communication, and problem solving. A teacher’s ability to develop these skills within students is imperative to a school attempting to break the cycle of poverty for their populations (Kane & Staiger, 2008; Wright, Horn, & Sanders, 1997). If a teacher’s generational low-income students cannot compete with their more affluent peers, they will not have access to the same career paths or post-secondary programs that will give them the means to escape the cycle of chronic poverty (Carnevale & Strohl, 2010; Roderick, Coca, & Nagaoka, 2008).

Furthermore, schools with low SES populations must prepare learners with a strong sense of agency and efficacy, who can adapt to the rapidly changing world economy and circumstances if they are going to be successful at breaking the cycle of poverty (Fadel, Bialik, & Trilling, 2015). Social cognitive theorist, Albert Bandura (2006) explained:

These transformative changes are placing a premium on the exercise of human agency to shape personal destinies and the national life of societies... Those of high self-efficacy influence the course of their occupational self-development, are receptive to innovations, and make their work life more productive and satisfying by restructuring their occupational roles and the processes by which their work is performed. (pp. 175 - 176)

Therefore, the scope and rapid pace at which economic, technological, and social changes are happening require that students have a level of agency to adapt quickly and profoundly.

Teachers' Roles in Developing Agency and Efficacy. Teachers working within these low SES populations are uniquely poised to facilitate the development of student agency and efficacy. As one of the most influential components of a school, teachers are frequently on the front lines of the war on poverty (Berliner, 2013). In contrast to the Coleman Report, Rivkin, Hanushek and Kain (2005) ascertained that the quality of teachers' instruction was more determinative of student achievement than family background. While teachers are important, urban, high poverty schools are not always able to facilitate this development to scale because they tend to have disproportionately higher numbers of ineffective and new teachers (Esch, Chang-Ross, Guha, Tiffany-Morales, & Shields, 2004; Lankford, Loeb, & Wykoff, 2002; Stuhlman & Pianta, 2009). These districts should therefore provide intentional, targeted support to equip teachers to be successful in these unique environments and circumstances to ensure the success of their students. Otherwise, the results can be alarming.

As the income gap has widened by approximately 40-50% over the last 25 years, the tremendous strains and pressures placed on high poverty schools and teachers in these environments have also increased (Reardon, 2011). Teacher burnout and turnover rates in high poverty, urban schools are unsettling (Allensworth, Ponisciak, & Mazzeo, 2009; Ingersoll, 2001; Ingersoll & Merrill, 2012; Johnson, Berg, & Donaldson, 2005; Marinell & Coca, 2013). In a study on Chicago Public Schools, Allensworth et al. (2009) found half of a typical schools' teacher work force turned over every five years, and only 33-37% of new teachers hired in 2002-2003 stayed at their schools for more than four years. Until poverty, and the social, physical, and financial burdens that come with it are eradicated, there will be a need for teachers to be

supported by their evaluation systems to remain and thrive in their positions to help students achieve (Ronfeldt, Loeb, & Wyckoff, 2013).

Supporting Teachers through Targeted Feedback. Because students from chronic poverty must rely on their teachers for much more than just academic knowledge, they need teachers who can persevere, adapt, and create learning environments where students engage in high quality experiences that develop the skills, knowledge, and beliefs required to move out of poverty (Downey, Von Hippel, & Hughes, 2008; Marinell & Coca, 2013). Moreover, it is the responsibility of district and school leaders to ensure teachers are supported and can maintain their focus on priorities that correlate most strongly to high student outcomes (Farnham et al., 2015).

Districts and school leaders should support teachers with intentional, targeted development through their evaluation system and feedback processes. Goldring et al., (2015) recognized the importance of evaluation systems as support processes and the value of measurement tools to emphasize areas as the priority outcomes for teachers and students. Specifically, the researchers identified that evaluative data can be used to guide “crucial conversations” between teachers and principals. These conversations are meant to identify both strengths and areas for improvement, and generate actionable feedback to help teachers better meet the needs of students in the classroom (Goldring et al., 2015).

Darling-Hammond, Wise and Pease (1983) recognized that effective teacher evaluation “requires consistent and shared views of the teaching-learning process and of the organizational context in which teacher evaluation takes place” (p. 286). To be effective in generationally high poverty settings, these evaluation systems and action plans need to prioritize and support the social and emotional needs of students as well as their academic. It is not enough that teachers

are proficient at teaching content to their students. Teachers in high poverty schools must first and foremost teach students; they must be caring adults who teach the whole child, including non-cognitive skills, not just academic content (Farnham, Fernando, Perigo, Brosman, & Tough, 2015; Farrington et al., 2012; Heckman & Kautz, 2012; Socol, 2014). High poverty, urban populations demand teachers be surrogate parents, mentors, social workers, counselors, cheerleaders, coaches, facilitators, and muses. Few of the roles on this list require the specific academic content knowledge teachers acquire in college, but they are incredibly important for reaching and teaching generationally impoverished children. These multifaceted relationships teachers build are especially important in high poverty settings considering students frequently report positive relationships as being integral to their engagement in school and overall academic success (Allen, Pianta, Gregory, Mikami, & Lun, 2011).

Evaluation and support systems of high poverty schools also need to recognize and prioritize the teaching of essential academic/career skills like collaboration, communication, complex problem solving, metacognition, self-regulation, time management, project planning, goal setting, academic resiliency, and how to accept and implement feedback and reflect on their learning and growth if students are going to be successful (Bouffard-Bouchard, Parent, & Larivée, 1991; Cleary & Callan, 2014; Duckworth, White, Matteucci, Shearer, & Gross, 2016; Duckworth & Yeager, 2015; Farrington et al., 2012; Heckman & Kautz, 2012; Lin-Siegler, Dweck, & Cohen, 2016; Mega, Ronconi, & De Beni, 2013; Mouratidis, Vansteenkiste, Michou, & Lens, 2013; Rutledge, Cohen-Vogel, Osborne-Lampkin, & Roberts, 2015; Tough, 2013). Teachers and students must be encouraged and supported to foster environments and cultures where academic pursuits take precedent over the personal obstacles and setbacks that come with living in poverty (Farrington, 2012; Ferguson & Danielson, 2012). Teachers in high poverty

schools must inspire and push students to have high aspirations, and model life-long learning (Ferguson & Danielson, 2014).

Intentionality in Designing Learning Experiences and Environments to Promote Agency and Efficacy. One such way to support long-term student growth is through the explicit teaching of non-cognitive skills, especially those that foster students' academic agency and efficacy beliefs. Non-cognitive skills (also known as soft-skills, deeper learning, social-emotional learning [SEL], self-regulated learning [SRL], and 21st century skills) are the attributes that comprise the essentials skills in addition to cognitive ability that contribute to aptitude and success. Efficacy, grit, growth-mindset, social well-being, self-regulation, social and emotional learning, self-awareness and metacognition, etc. all fall under the non-cognitive umbrella (Duckworth & Yeager, 2015; Heckman & Kautz, 2012; West, Kraft, Finn, Martin, Duckworth, Gabrieli & Gabrieli, 2016). Evans and Rosenbaum (2008) observed that academic disparities in affluent and disadvantaged students are correlated to their level of non-cognitive skills. Although non-cognitive skills have been shown to be relationally important to student achievement, these cannot be taught in the same ways as academic knowledge (Farnham et al., 2015; Tough, 2015).

After some backlash about his popular book, *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character*, Tough (2015) called for a shift in the rhetoric from teaching non-cognitive skills through isolated activities to intentionally creating classroom environments that demand, explicitly teach, and foster these skills daily in the context of academics.¹ He maintained that to create more effective students, teachers and school leaders must change the

¹ Several articles have highlighted the concerns researchers and academics have with the “grit” narrative as it pertains to students in poverty or from disadvantaged backgrounds, including a peer reviewed article by Ris (2015) and an OpEd article from Herold (2015).

learning environment (Tough, 2015). It is not enough to have isolated lessons on perseverance, goal setting, or problem solving, or to even to have students engage in these activities periodically. A culture fostering students' SEL and academic efficacy beliefs must be embedded in the fiber of the classroom and school to have a substantial influence on students' achievement and agency (Farnham et al., 2015; Farrington et al., 2012; Tough, 2015). Cultivating such a culture has implications for teacher evaluation systems, which will be discussed in later sections.

Furthermore, SEL is already utilized for some level of school accountability in a handful of California school districts. Eight large urban districts, known as the CORE districts and comprising over a 20% of the state's student population, are currently undergoing a massive study on their adopted accountability framework that incorporates academic, SEL, and other culture and climate measurements in their school quality index (Krachman, Arnold & Larocca, 2016). Their partner organization, Transforming Education (2016), identified specific Mindsets, Essential Skills, and Habits (MESH), which demonstrated a high correlation with students' success in college, career, and life. Their priority SEL competencies are growth mindset, self-efficacy, self-management, and social awareness. Their findings and research process will be explored further in the literature review.

The Problem

The problem is students from low SES backgrounds are not achieving commensurate with their high SES counterparts. They need to develop both academic prowess and the SEL behaviors to support their academic achievement. Understanding the factors that may influence this discrepancy should aide in the development of targeted interventions to support teachers and their students. This study uses the Social Cognitive Theory to explore those factors, including student factors (self-efficacy, demographic information), environment factors (teachers'

perceptions of their students, behaviors, classroom cultures, school/classroom demographics like looping), and how these factors interact with each other to influence student behaviors (including SEL and academic achievement).

Environment Factors. There is a lack of literature seeking to define features of agentic classroom cultures that can be utilized to create effective evaluation criteria related to the teacher's role in developing the agency and efficacy of their students. Therefore, there is a need to ascertain whether students' agentic related behaviors can be observed in classroom contexts, and whether these data can be used reliably to measure teacher effectiveness.

Student Factors. Typically, data regarding student efficacy are collected via survey. However, several researchers have called for new methods for studying academic efficacy levels beyond surveys (Duckworth & Yeager, 2015; Stecher & Hamilton, 2014). Classroom observations and informal discussions with students on their efficacy and agency in the context of learning can allow for the documentation of these defining features that can lead to further support of teachers and growth of students.

Interactions Among Factors. Past efficacy studies have focused primarily on individual student efficacy or collective teacher efficacy and their positive relationship with student achievement (Farrington et al., 2012; Foster, 2015; Goddard, 2001; Mega, Ronconi, & De Beni, 2014; Multon, Brown, & Lent, 1991; Schunk, 1989; West, 2014; Zahodne, Nowinski, Gershon, & Manly, 2015; Zimmerman, 1999). However, there is a void in the literature related to the levels of students' efficacy in the context of the classroom and the teacher's role in cultivating agency in that context.

Purpose of the Study

The purpose of this study was to explore factors related to Social Cognitive Theory that influence achievement and student success using concurrent triangulation mixed methods (Creswell, 2014). This study identified the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in high poverty classrooms. It is my hope these findings could be used to better support teachers and evaluation systems in the future. In the first quantitative phase of the study, the research questions explored the relationship between student achievement, perceived student agency and academic efficacy in classrooms, and the students' levels of perceived teacher effectiveness. The independent variables in this phase were demographics, perceived self-efficacy, self-management, and social awareness, student perception of teacher, and teacher perception of student as measured by student and teacher surveys. The dependent variables were academic achievement in ELA and mathematics as measured by Acuity (McGraw Hill) assessments. The purpose of the concurrent, qualitative Phase II of the study was to identify characteristics of classrooms and practices of teachers that establish agentic behaviors and academic self-efficacy beliefs of students.

Research Questions

The research questions this study answered were:

Quantitative Phase:

1. What are the relationships among students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness?

Qualitative Phase:

2. What are the observable behaviors and norms of students in classrooms with varying levels of cultures of agency and self-efficacy (CASE)?
3. What relationship do students' agentic and efficacious classroom behaviors have with their perceived self-efficacy levels and academic achievement?

Overall:

4. What are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?

Significance

When No Child Left Behind (NCLB) was replaced in 2015 with the Every Student Succeeds Act (ESSA), new methods and measures for school, teacher, and student evaluation and teacher effectiveness became policy possibilities. Darling-Hammond, Bae, Cook-Harvey, Lam, Mercer, Podolsky, and Leisy Stosich (2016) noted the complex shift in this process necessitated by ESSA:

A system that focuses on the whole child and the whole school requires a more comprehensive set of indicators that measure the range of skills and competencies students need to be successful upon graduating from high school. These include the mastery of core academic content; the ability to think critically, collaborate, and communicate; the development of academic mindsets; and the capacity for independent learning. The resources and conditions that support students' opportunities to learn must also be included (p. 2)... Thus, states should thoughtfully consider how each element of their accountability system creates incentives and opportunities to move school practices forward in ways that better ensure that all students are successful in their learning and

their lives beyond school. Part of this process is identifying what kinds of conditions and practices have been found to lead to better outcomes for students. Those that have greater influence on student success should have special consideration as potential indicators. Furthermore, measures of learning should seek to capture the aspects of student performance that have greatest traction for later success...(p. 4).

While the usefulness of this study may not extend to state level accountability measures, it could serve to improve and measure teacher effectiveness at fostering environments that impact student attitudes and aptitudes for academic learning. In his analysis on California's CORE districts measuring students' growth mindset, self-efficacy, self-management, and social awareness, West (2016) noted that the deeper purpose of evaluation and accountability systems is to help teachers understand what is important and what will facilitate student growth. The significance of this study is the potential for the eventual measurement and inclusion of indicators of agency and academic efficacy in teacher evaluation systems for improving educational practice, and the non-cognitive and academic outcomes for students.

Delimitations

This study included teachers and students in grades 3-5 in two public charter schools located in the urban core of Indianapolis. They were chosen because over 95% of their populations are impoverished, as measured by free and reduced price lunch calculations, to ensure any variations in what I am observing are not due to differences in SES. The demographic complexities of the two schools also met or exceeded those of the local traditional public schools. They were two public charter schools within the same network, with a history of student-centered teacher evaluation. One was a growing K-5 elementary and the other was a combined K-12 school. All the teachers at the schools had had varying degrees of training on

SEL particularly in the areas of growth mindset and self-efficacy, and all students were assigned to teachers and classrooms in an equally stratified manner in terms of English Language Learner (ELL) and Special Education (SpEd) designations, and academic and behavioral abilities.

Definition of Terms

These key terms are defined within the context of this study:

Academic efficacy. Academic efficacy is students' beliefs about their academic capabilities, and overall ability to apply their knowledge and skills to produce new learning, acquire a new set of skills, or complete a learning task (Bandura, 1997; Usher, 2009).

Agency. Bandura's (2006) definition of agency is the act of a person intentionally influencing and contributing to his/her functioning and life circumstances.

Agentic and efficacious behaviors. The visible, outward behaviors students display when they are struggling with a task, concept, behavior, or class that help bridge the gap between struggle or failure and success (Axelrod, Bellini, & Markoff, 2016)

Collective efficacy. When efficacy beliefs are expanded to include those of the group it is called collective efficacy (Bandura, 1997). Rather than "the sum of the efficacy beliefs of individuals... it is an emergent group-level attribute that is the product of coordinative and interactive dynamics" (Bandura, 1997, p. 7). Most educational research in the field of collective efficacy has focused on collective teacher efficacy as opposed to students'.

Evaluation. Evaluation is a feedback process designed to help teachers improve while measuring teacher competency and student outcomes (Marzano & Toth, 2013).

MESH. This is an acronym for Transforming Education's (2016) Mindsets, Essential Skills, and Habits that have been shown to have a strong relationship with students' success in college, career, and life.

Growth mindset. Dweck (2006) described a growth mindset as the belief that a person can change and intentionally cultivate who they are through effort and application of experiences. Specifically, people who have a tendency to stretch themselves, love learning, persevere, and even thrive when things are challenging or difficult are characterized by a growth mindset.

Looping. Looping is the act of assigning a cohort of students to the same teacher for two or more consecutive years (Cistone & Shneyderman, 2004).

Non-cognitive skills. The appropriateness of the term non-cognitive skills has been debated and discussed by policy makers, researchers, and educators (Duckworth & Yeager, 2015; Farrington et al., 2012). Terms such as social and emotional learning (SEL) (Durlak, Domitrovich, Weissberg, & Gullotta, 2015), 21st century skills (Soland, Hamilton, & Stecher, 2013), soft skills (Heckman & Kautz, 2012), and non-cognitive functions (Farrington et al., 2012) have also been used to describe the behaviors, attitudes, beliefs, strategies, mindsets, and skills students have beyond cognition that contribute to the academic achievement of students. These cannot be measured by typical IQ tests and are not currently evaluated on standardized assessments, but have been shown to have a positive correlation to academic outcomes.

Reliability. The consistency of the judgments of observers based on the data collected during evaluation walk-throughs and observations (Gullickson, 2009).

Rubric. The measurement tool used most commonly to measure teacher effectiveness is a rubric. A rubric is a set of measures that includes descriptions of levels or dimensions of quality on specific criteria (Brookhart, 2013).

Self-efficacy. The perception of self-efficacy is a person's belief that they have the ability to act in a way that will produce desired outcomes. These actions may involve managing or changing their motivation, thoughts, affect, actions and environment (Bandura, 1997).

Self-management or Self-regulation. This is the ability of a person to effectively control their thoughts and actions that allow them to work toward and meet goals (Transforming Education, 2016).

Social and emotional learning (SEL). The behaviors, attitudes, beliefs, strategies, mindsets, and skills students have beyond cognition that contribute to their academic achievement (Durlak, Domitrovich, Weissberg, & Gullotta, 2015).

Social awareness. Social awareness is understanding norms of behavior, particularly for communicating and working with peers to solve problems and conflicts, and recognizing the resources and support systems available to further personal learning and growth (Transforming Education, 2016).

Value added measures or models (VAM). Value added measures or models in teacher evaluation involve measuring student outcomes relative to prior performance and growth over the year relative to peer groups to determine the effectiveness of teachers' instruction (Harris, 2011).

Organization of the Study

This chapter presented background information regarding this study and provided an overview for the forthcoming research. Chapter Two contains my conceptual framework and a comprehensive literature review on the topics of teacher evaluation and social cognitive theory (SCT). Chapter Three outlines the study's research design and methods. Chapter Four contains analysis and discussion of the findings, while Chapter Five contains a summary, conclusions, and

recommendations for further research. Finally, the resources and appendices conclude this study.

CHAPTER 2

LITERATURE REVIEW

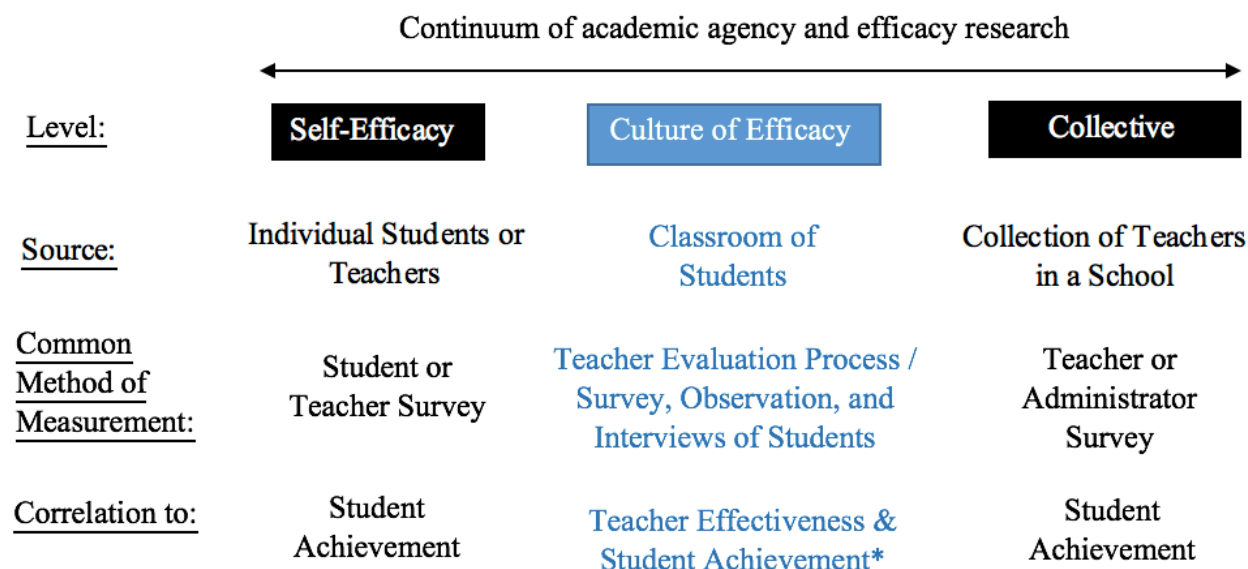
A teacher's ability to create conditions in their urban, high poverty classrooms that cultivate students' agency and self-efficacy has been the subject of few research projects despite the number of studies that have demonstrated the strong correlation between students' levels of self-efficacy and their academic achievement (Farrington et al., 2012; Foster, 2015; Mega, Ronconi, & De Beni, 2014; Multon, Brown, & Lent, 1991; Schunk, 1989; West, 2014; Zahodne, Nowinski, Gershon, & Manly, 2015; Zimmerman, 1999). Furthermore, researchers have noted that to be effective, social and emotional learning (SEL) and academic self-efficacy must be embedded in the fiber of the classroom and school to have a substantial influence on students' achievement and agency (Farnham et al., 2015; Farrington et al., 2012; Tough, 2015). This chapter examines past research, beginning with an exploration of the conceptual framework created for this study.

I used an original conceptual framework supported by two research-based theoretical frameworks when conducting this analysis. The classroom Culture of Agency and Self-Efficacy (CASE) conceptual framework developed for this study was used to synthesize social cognitive theory (SCT) and teacher evaluation, and justify and explain the classroom focus of this research. Gullickson and Howard's (2009) *Personnel Evaluation Standards*, specifically the standards on Constructive Orientation, Valid Judgments, and Reliable Information, were used to guide the orientation and measurement determination process for teacher effectiveness. SCT was used to guide the lens for the specific indicators during observations, informal discussions with students, and analysis of teacher effectiveness.

Chapter 2 is divided into five main sections. The first section outlines the conceptual framework, CASE, and serves to simply allow the reader to understand the relationship between teacher evaluation and SCT explored in the second and third sections. The second and third sections begin with a description of the evaluation and SCT frameworks and then transition to the review of the literature for each. I conclude each section with a synthesis of the instructional and evaluative implications for the classroom and/or my study. A summary concludes the chapter.

Conceptual Framework Preview – The Culture of Agency and Efficacy

My CASE conceptual framework maintains that between Bandura's (1997) measurement of the individual self-efficacy of students, and the collective teacher efficacy of the organization, there exists a measurable level of agency and self-efficacy cultivated by the teacher and students in the classroom environment. Furthermore, the evaluation of the agentic classroom culture can be utilized to measure a teacher's effectiveness, similar to the way 'respectful culture' is currently utilized by several prominent evaluation systems as an indicator of effectiveness (Danielson, 2013; National Institute for Excellence in Teaching, 1999). This agentic and self-efficacious culture includes norms, beliefs, behaviors, and language that build and indicate the presence of student agency and academic efficacy. Figure 1 illustrates this framework on the continuum of academic efficacy research, and an illustration of the focus, measurement, and outcome correlations of each.



* This is unproven. Potential for confirmation with this and subsequent studies.

Figure 1. The relationship of the Culture of Agency and Self-Efficacy (CASE) conceptual framework to self and collective efficacy.

A student's agency and efficacy beliefs are affected by their behavior, the environment, and other personal factors including biology, cognition, and affective events (Bandura, 1997). Teachers have enormous control and influence in their classrooms (Kane & Staiger, 2008; Wright, Horn, & Sanders, 1997). To overcome negative influences in factors outside of direct control, a teacher in a high poverty, urban school must teach students how to regulate their behavior, manipulate their environments to be successful, and support development of positive personal factors. The two areas identified by Bandura (1997) that teachers can help students control are their own behavior and cognition. This influence must be so powerful and impactful that it can withstand and overcome the significant environmental influences outside of school that work to counter low SES students' efficacy beliefs (Bandura, 1997).

Many students from chronic poverty bring a set of, often underappreciated, skills to the classroom, which contribute to their level of efficacy and agency outside of academia (Berliner, 2013; Socol, 2014). Many of these students are resilient, adapting and surviving circumstances and life events that would intimidate many adults. They take on adult roles and responsibilities at home and act as caregivers for younger siblings. They are resourceful and creative in overcoming situations, and have an ability to conceive of opportunities where others would see none (Mullainathan & Shafir, 2013; Socol, 2014). These traits may facilitate academic agency and efficacy if these students can transfer their resilience and adaptability to their academic pursuits (Socol, 2014). It is important for teachers to understand this, to foster behaviors that help with transfer of student understanding and non-cognitive skills to academic contexts, and deliberately establish learning environments where efficacy is cultivated and fostered (Tough, 2015).

Several studies have ascertained that self-efficacy can be cultivated through the intentional construction of learning environments to help students effectively process experiences (Farnham et al., 2015; Phan, 2013; Schunk & Rice, 1987). Findings from Schunk and Rice (1987) suggested teachers should not only create lessons and activities that teach multiple strategies to approach problem solving and interpret feedback, but also explore explicitly why they are effective and under which circumstances. The researchers support that this is an integral part of academic success for students (Schunk & Rice, 1987). Teaching students to recognize why a strategy is useful and when to apply it can contribute to their growing sense of agency and self-efficacy by giving them a degree of control over their learning (Schunk & Rice, 1987). Teachers can then look for these student behaviors to help inform

whether their classroom environment and learning tasks are contributing to the cultivation of efficacy in their classrooms.

If teachers can explicitly establish learning environments, teach, and support the means that allow students to accomplish tasks and act in ways that cultivate success, students will feel more in control of their own development and increase agency outcomes and efficacy beliefs (Bandura, 1997; Schunk & Rice, 1987). Students belong to a wider community of learners that make up the classroom (Putney & Broughton, 2011). Therefore, for evaluation and support purposes, teachers should receive feedback on the collective culture and tendencies of the students' behaviors in their classrooms, making a culture of efficacy a potentially important indicator of teacher effectiveness.

A rare study on classroom level collective student efficacy is an ethnography of one classroom over the course of four years. Located in a school where 85% of the population was of low SES, Putney and Broughton (2011) described their observations of the teacher's role and her ability to develop a strong sense of what they called "collective classroom efficacy", year after year. Specifically, the researchers were trying to understand how the teacher, Ms. Falls, fostered a sense of collective responsibility, and whether that responsibility and the sense of belonging that characterized the classroom related to collective efficacy (Putney & Broughton, 2011). The researchers collected examples of individual and collective efficacy as described by Bandura (1997). They found that Ms. Falls gradually released responsibility of the classroom functions over to the students and acted as both supporter and facilitator for learning opportunities. As Bandura (1997) described, she acted as the community organizer (Putney & Broughton, 2011). Students self-assessed and received feedback on academic work. Ms. Falls used a "work ethic rubric," to encourage students to take academic and intellectual risks, lead

class conversations, critique each other's work, give feedback to one another, and participate in activities that developed their life skills throughout the year (Putney & Broughton, 2011, p. 98). In general, Putney and Broughton (2011) recommended that teachers who want to serve as community organizers establish and encourage "informed [academic] risk taking by (a) creating a sense of belonging, (b) setting and working toward personal and academic goal attainment, (c) taking responsibility for self and others' learning, and (d) believing in individual and collective capabilities" (p. 103).

Putney and Broughton (2011) recognized that their research was limited in scope to one teacher for describing the classroom transformation into a very specific learning community. They encouraged others to conduct further research on the relationship between collective classroom efficacy and student achievement, and on the role of the teacher as a community organizer. This research was a seminal piece for my dissertation because it called for the subject and type of research I sought to conduct. Further discussion on the role of teacher as community organizer is contained in Chapter 5.

Due to the limited research on evaluating teachers on the agentic and academically efficacious behaviors of their students and cultures of their classrooms, the effective organization of the literature review was challenging. To help with flow and delineation of ideas, the remaining sections of this chapter are divided into teacher evaluation and Bandura's (1986) social cognitive theory, including agency and self-efficacy. Throughout each section, I also explore the implications for application of both teacher evaluation and social cognitive theory in the context of the classroom under the CASE framework.

A Review of Teacher Evaluation Literature

The concept of teacher evaluation is over a century old, yet gained prominence in recent years due primarily to the U.S Department of Education's distribution of \$4.3 billion to states in 2009 via the Race to the Top (RTTT) grant (Sartain, Stoelinga, & Krone, 2010). For perspective, only 15 states required any type of annual teacher evaluation in 2009; by 2013, 45 states had adopted evaluation policy that specifically required some form of evaluation for teachers and 35 use evaluations to determine effectiveness (Doherty & Jacobs, 2015).

Per the Highly Qualified Teacher provision of No Child Left Behind (2001), all teachers needed a minimum of a bachelor's degree and a state certification or licensure with proven knowledge of subject(s) they teach (US Department of Education, 2004). However, highly qualified did not necessarily equate to highly effective. At the root of RTTT (2009) and the evaluation movement was the desire by policy makers and education leaders to differentiate teacher quality in a system that largely ignored quality. Rather, teachers were compensated based on years of experience, despite multiple bodies of research that concluded there was not a relationship between teacher experience and students' achievement (i.e., Darling-Hammond & Youngs, 2002; Jacob & Lefgren, 2008). In 2009, no states used evidence of teachers' effectiveness to make tenure decisions. As of 2015, 23 states used teacher performance for such decisions (Doherty & Jacobs, 2015).

Between the passage of NCLB (2001) and the introduction of RTTT (2009), researchers set out to quantify the impact of NCLB's Highly Qualified Teacher provision on student learning for the primary purpose of defining quality and sorting teachers. Many established that the teacher evaluation process, used strategically with a combination of support and feedback, helps clarify, prioritize, and improve instructional practices. As the Reform Support Network (2015)

sponsored by the U.S. Department of Education noted, “The ultimate goal of teacher evaluation systems is to improve the quality of instruction by clarifying expectations for effective teaching and helping teachers meet those expectations through high-quality feedback and support” (p. 1). The following paragraphs explore this research, including the process for improving instruction, the measurement tools or rubrics, practical measurements, and the relationships between evaluative measures and processes and student achievement. This section concludes with an analysis of the implications for this study.

Process for Improvement

Although the scope of this study did not extend to providing support and feedback to teachers to improve their instruction, it should be the purpose of any evaluation tool or rubric to help teachers grow as practitioners. Constructive evaluation processes and rubrics should support the success of all students, teachers, and schools, and propel them towards meeting organizational goals (Gullickson & Howard, 2009). The rubric and/or evaluative process should not be used to simply identify weaknesses, control or mandate actions to teachers, dismiss instructional personnel, or be used in any way that would discourage educators. Rather, it is my hope that beyond the research of this study, the observation tool I developed could be studied further and an iteration be eventually used to improve outcomes for students and “provide educators with information and professional feedback that build their professional self-knowledge, increase their enthusiasm, and enhance their efficacy as practitioners” (Gullickson & Howard, 2009, p. 70). Therefore, I drafted this tool with these purposes in mind.

One focus of teacher evaluation research in recent years has been on the systematic and intentional support to improve the quality of instruction through multiple observations and quality feedback. Researchers suggest teacher evaluation should inform professional

development and facilitate growth (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Hill & Grossman, 2013; Kane, McCaffrey, Miller, & Staiger, 2013; Papay, 2012; Ritter & Barnett, 2016; Weisber, Sexton, Mulhern, & Keeling, 2009). Weisber et al. (2009) concluded districts should adopt evaluation processes that differentiate teacher performance and provide targeted support for teachers; the system should have performance standards rooted in student outcomes, and require regular monitoring and feedback cycles; and evaluators should receive training to effectively measure teacher performance and provide quality feedback to teachers on rigorous performance standards. Darling-Hammond et al. (2012) found that multiple and frequent observations with evaluators using multiple sources of evidence were features of the most successful systems.

Weaving the evaluation and support structures together into one system could be highly effective way to ensure the support, quality feedback and measurement goals are seamless and effective (Ritter & Barnett, 2016). Specific programs like the Peer Assistance Review (PAR) program in California pair coaches with teachers at all levels of experience and proficiency. Coaches conduct frequent classrooms observations as a part of the evaluative process and provide feedback and targeted professional development to teachers in an effort to improve their craft and student outcomes (Goldstein, & Noguera, 2006).

The utilization of coaches to both support and evaluate teachers has been shown as a valid means for improving teacher effectiveness. TAP, an evaluation system relying on a series of evaluations conducted by master teachers and mentor teachers on career teachers, has been lauded as effective improvement system (Darling-Hammond et al., 2012; Ritter & Barnett, 2016). Hill and Grossman (2013) recommended that coaches and department chairs conduct observations and provide specialized feedback as a part of the evaluation system because of their

expertise with the subject matter, student population, and specialized skills required for success. Other studies have demonstrated that teachers who are provided individualized support and feedback are more effective as measured by student learning and behavior outcomes (Biancarosa, Bryk, & Dexter, 2010; Flynn, Lissy, Alicea, Tazartes, & McKay, 2016; Vanderburg & Stephens, 2010). Specifically, Vanderburg and Stephens (2010) ascertained that coaching helped teachers innovate in the classroom and improve their overall sense of agency, which led to an increase in the level of agency and achievement in students.

Measurement Tools

The measurement tools used to define quality and evaluate teachers are typically rubrics. According to the Reform Support Network (2015), rubrics should be coherent and aligned with a state's standards for instruction, be concise and brief, clearly describe teacher and student behavior, and contain focused indicators that "directly relate to student outcomes" (p. 2). Generally, these rubrics are categorized as holistic or analytic (Arter & McTighe, 2001; Davis, 2016). Holistic rubrics describe the general quality of the overall body of work and use general language such as 'demonstrates understanding' or 'few mistakes' or 'all requirements are included.' In contrast, an analytical rubric articulates specifically what each level of criteria entails and the language and depth of knowledge needed to demonstrate competency. Analytical rubrics can also be used to describe the developmental process a person would go through in acquiring competency (Arter & McTighe, 2001; Davis, 2016).

The process for developing an analytical rubric has been documented by several leading researchers and authors (Allen & Tanner, 2006; Arter & McTighe, 2001). Notably, Arter and McTighe (2001) recommended a developer begin with a review of the literature, progress into an examination of student work, sorting the work by quality, and then developing criteria to

describe the quality of each set of samples. Finally, they recommend the developer continue to find student samples for continued scoring and refinement of the criteria and validation of the rubric.

Other researchers have studied the focus of evaluation instruments and what they measure. In their research on 16 different observation and evaluation systems from around the world, Martinez, Taut, and Schaff (2016) found American systems tended to focus on the technical and procedural components of instruction (for example, teacher's ability to manage the classroom, track data, and ask higher level questions) whereas Japan and Singapore's systems inferred teacher competency based on students' academic growth and holistic aptitudes (for example, the students' values and self-confidence). Japan and Singapore are notable because of their status as two of the world's highest performing education systems (Steiner, 2010). In Singapore, observers collect evidence of the nurturing of the whole child through intensive observation, which is then holistically considered for evaluative purposes. Principals in Japan are also encouraged to holistically use evidence from observations conducted by administrators and other senior teachers to provide feedback and support, and develop individualized plans for teachers (Martinez et al., 2016).

Similarly, Ovando (2001) maintained evaluation and support systems used in American schools, including the measurement tools, should be learner-centered to drive student achievement and teacher improvement. Utilizing the walk-through process, these quick observations allow time for task-specific feedback and that create "opportunities for growth, feedback, learner-centered dialogue, a holistic perspective, and teacher self-evaluation" (Ovando, 2001, p. 228). Recent evaluation frameworks developed for U.S. schools employ feedback and self-evaluation processes and measurement tools that focus more on the learner actions and

outcomes rather than the teacher's. Revisions made to Danielson's 2007 framework included Critical Attributes and Possible Example criteria that helped frame observer feedback in the context of both student and teacher behaviors (Danielson, 2013). The New Teacher Project's (TNTP) (2015) Core Teaching Rubric used the word "teacher" or its derivative 42 times, whereas the word "student" appeared over 200% more times, for a total of 140 instances. Because I was interested in students' agentic and efficacious behaviors, I also framed the evaluative criteria in my measurement tool in the context of the degree of "ownership" of the behaviors, or the extent to which students were independently or dependently engaging in the classroom behaviors versus the teacher.

Practical Measures

Researchers in the last few years have been promoting the use of "practical measurements" to improve practice as opposed to accountability measures or theory validation (Duckworth & Yeager, 2015; Yeager, Bryk, Muhich, Hausman, & Morales, 2013). Duckworth and Yeager (2015) described practical measurements in this manner:

[Practical measurements] are administrable in the web of daily instruction, they can be quickly reported on and communicated to practitioners, and they have direct relation to causes of student underperformance that are the explicit target of improvement efforts. They allow people to learn rapidly from practice. This means that the measures should be brief, easily collected, and also contextually appropriate. Practical measure should be sensitive to short-term changes and provide short-term feedback on progress that has or has not been made in improving personal qualities (p. 245).

It was my intent to utilize a very practical data collection process and to develop a tool that could allow the measurement of student behaviors in the context of classroom culture quickly to

accurately capture the throughputs and outputs of effective instruction and environmental norming rather than inputs. My goal was to create a tool that could eventually be useful to the field as well as to academia.

Evaluation and Student Achievement

Recently, teacher evaluation research has examined the relationship between evaluative measurements and student achievement. The Bill and Melinda Gates Foundation had two major aims upon launching the largest study to date on teacher effectiveness in 2009 in six school districts across the nation. The Measures of Effective Teaching (MET) project was predicated on the beliefs that, “First, a teacher’s evaluation should depend to a significant extent on his/her students’ achievement gains; second, any additional components of the evaluation (e.g., classroom observations) should be valid predictors of student achievement gains” (Bill & Melinda Gates Foundation, 2010, pp. 4-5). Over 20,000 videos were collected from classrooms in Charlotte-Mecklenburg Schools (NC), Dallas Independent School District (TX), Denver Public Schools (CO), Hillsborough County Public Schools (FL), Memphis City Schools (TN), and the New York City Department of Education (NY). Each video was analyzed using several scales including Danielson’s (2007) Framework for Teaching and the Classroom Assessment Scoring System (CLASS) (Pianta, La Paro, & Hamre, 2008). The findings are discussed in the following paragraphs.

The process. Strong (2011) created the Rapid Assessment of Teacher Effectiveness (RATE) to evaluate videos using CLASS. Classrooms that trended above and below average in effectiveness as defined by student achievement scores were chosen. Specifically, RATE was created to predict the effectiveness of teachers based on observations of their teaching as compared to the achievement of their students (Gargani & Strong, 2014). After training

observers for only four hours, Strong (2011) established that seven indicators, all in the instruction domain on the CLASS, accurately identified teachers as effective or ineffective as compared to their student achievement data. These indicators were ‘clear expression of the lesson objective, integrating students’ prior knowledge, using opportunities to go beyond the current lesson, use of more than one delivery mechanism or modality, using multiple examples, giving feedback about process, and asking “how” and “why” questions’ (Strong, 2011, p. 98).

Strong (2011) cautioned that the study was small, but could be used as an efficient way to conduct future teacher evaluations including those conducted on the MET videos. Furthermore, Strong’s (2011) results were also significant regarding identifying sources of agency and efficacy in a classroom. Having a clear goal, the use of multiple examples, and feedback about the process were all described as sources of efficacy via enactive mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states by Bandura (1997). These sources are explored in later sections.

Random vs. non-random assignment. During the second year of the MET project the districts were required to randomly assign students to teachers and did so with varying degrees of fidelity. After pulling data for only those classrooms where students were randomly assigned, Kane et al. (2013) determined that those teachers demonstrated a .955 coefficient of effectiveness on student achievement, with a standard error of .123 when mathematics and ELA achievement results were combined (p. 25). The researchers concluded that effective teachers did affect student achievement particularly when they were randomly assigned students. This suggests random, heterogeneous student groups are an important variable to reliable value added models (VAMs) (Kane et al., 2013).

However, random assignment to classrooms rarely occurs in practical applications and authentic school contexts. Deliberate assignment of students to particular teachers' classrooms and for specific purposes are a reality in many districts and has been shown to have a strong relationship to evaluation scores and VAMs (Briggs & Domingue, 2011; Darling-Hammond et al., 2012; Kalogrides, Loeb & Beteille, 2013; Steinberg & Garret, 2016). Steinberg and Garret (2016) noted that the MET study had access to data from classrooms where student assignments were not random (51% of ELA and 61% of mathematics classrooms) and set out to compare the nonrandom classrooms to the findings of Kane et al. (2013). They ascertained that ELA teachers who were assigned the highest achieving students were more than twice as likely to be rated in the top quintile than their peers assigned lower achieving students. Mathematics teachers were six times as likely to be rated in the top quartile under the same conditions (Steinberg & Garret, 2016). In essence, Steinberg and Garret (2016) determined that effectiveness ratings of teachers with non-randomly assigned students were largely determined by students' prior performance.

These results confirmed a 2011 study by Briggs and Domingue (2011), who conducted similar research on Los Angeles Unified School District, and determined elementary students were intentionally sorted resulting in over half of ELA teachers and almost 40% of mathematics teachers rating differently when alternative models of VAMs were used to account for the non-randomness of classroom assignments. Kalogrides et al. (2013) analyzed data from classrooms in Miami-Dade County Public Schools and found that less experienced teachers, minorities, women, and teachers from less prestigious undergraduate programs were more likely to be intentionally assigned lower achieving students than their counterparts. As a result, they were more likely to be rated less effective, an indicator that nonrandom sorting can affect evaluation results (Darling-Hammond et al., 2012). This is a significant finding particularly for urban, high

poverty settings where the goal should be to assign teachers who can transform the lives of low SES students, not to replicate the academic factors that contribute to poverty.

Steinberg and Garret (2016) also established that non-random sorting influenced the correlation between evaluation indicators on measurement tools and the ratings of teachers with higher achieving students. For ELA teachers, the measurement indicators showing a high correlation with student achievement were *communicating with students* and *engaging students in learning*. For mathematics teachers, the correlation was with the indicator of *establishing a culture of learning in the classroom*. The relationships could be symptoms of the general significant lack of student dialogue that typically occurs in classrooms with lower achieving students. For example, the emphasis in low performing mathematics classrooms tends toward rote skill building and procedural algorithmic processes, whereas rich dialogue about numeracy is reserved for higher achieving students (Steinberg & Garrett, 2016). Interestingly, the researchers questioned whether the curricular and instructional shifts required under the Common Core State Standards would eventually move instruction to become more constructivist in nature rendering these correlations negligible. Steinberg and Garrett (2016) concluded that with nonrandom classrooms, the relationship between evaluation scores on specific indicators of measurement tools might be manufactured by the perceived characteristics of the students. This had implications for the methods of my study as discussed in Chapter 3.

Implications for This Study

Researchers are calling for evaluation processes, systems, and policy that can give a more complete and accurate picture of classroom learning especially in the context of ESSA (Darling-Hammond et al., 2012; Steinberg & Garret, 2016). I aspired to bring CASE to the teacher evaluation and support conversation, particularly for high poverty schools where agency and

self-efficacy in academics could be a powerful and effective means for student success and empowerment. My research process was a holistic, multi-step analysis of the data to draw conclusions about teacher effectiveness regarding CASE, and is detailed further in Chapter 3. It was also my intent for the tool I developed to be useful for providing context to feedback and support of teachers outside of the scope of this study, and was therefore constructed with this in mind.

Social Cognitive Theory – Agency and Efficacy

Social Cognitive Theory (SCT) was developed by Albert Bandura (1986) and was designed to describe the relationship between knowledge that a person acquires, the observations they make during their experiences, and their interactions with others and the environment. This reciprocal relationship between the person, behaviors, and the environment then correlate to sense of self. The concepts of agency and efficacy are under the umbrella of SCT.

The agentic perspective of SCT consists of four properties as described by Bandura (2006): intentionality, forethought, self-reactiveness, and self-reflectiveness. Each of these properties is defined as follows:

1. Intentionality - Bandura (2006) explained that people intentionally self-organize action plans and strategies for realizing self-interested goals. This is applicable to both self and collective efficacy constructs.
2. Forethought – People exhibit forethought thorough visualizing outcomes and the future, and using that as a guide and motivator for present decisions and actions.
3. Self-reaction - People self-regulate and control their reactions to stimuli and the environment to set and regulate their course of actions and motivate themselves.

4. Self-reflection - People also examine their thoughts and actions and adjust based upon these reflections. Bandura (2006) noted, “The metacognitive capability to reflect upon oneself and the adequacy of one’s thoughts and actions is the most distinctly human core property of agency” (p. 165).

To become agentic, individuals must frequently be placed in supportive situations where they can refine and practice their performance and decision making to make actions automatic. This example is similar to an athlete practicing their mechanics in various practice situations so their reflexes are automatic, producing the desired outcome in game-time scenarios. Bandura (2006) noted that students must be accorded a level of license to explore, manipulate, and influence their learning environment as well:

To make their way successfully through a complex world full of challenges and hazards, people have to make sound judgments about their capabilities, anticipate the probable effects of different events and courses of action, size up socio-structural opportunities and constraints, and regulate their behavior accordingly (p.168).

Furthermore, agency is not simply a private affair. Social experiences and the environment influence agentic perspectives and automaticity. Dialogue, interactions, expectations, assistance from parents, family, community members, teachers, peers, and the context of circumstances all play a role. Therefore, I used this framework to make the assumption that all students would benefit from learning in environments where agency is explicitly taught as a skill, frequently utilized by students in intentional situations, and embedded into the culture of the classroom in order to shape the skills, brain function and automaticity of their agentic beliefs and actions.

The other construct extracted from SCT used in this study is self-efficacy. Bandura (1997) described a person's efficacy as the central and most pervasive mechanism of their agency. A person's belief about whether they are capable of achieving something - that they can act in ways or produce behaviors that can result in achievement of determined goals or outcomes – is the foundation of agency. The efficacy framework has been applied in the fields of education, physical health, mental health, athletics, politics, and organizations to name a few. This theory indicates that intentional actions people put forth to reach their goals is derived from their sense of efficacy (Bandura, 2006).

Sources of Self-Efficacy

Bandura (1997) identified four sources of self-efficacy information that cause the reinforcement or revision of efficacy beliefs. These are enactive mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states. Moreover, the integration process of this information into the psyche is influenced by the perceived value of the information and the degree of the interrelatedness between prior efficacy beliefs and the new construct (Bandura, 1997). According to Bandura (1997), "Efficacy beliefs contribute to the quality of human functioning in diverse ways. They do so by enlisting cognitive, motivation, affective, and decisional processes through which accomplishments are realized" (p. 115).

Enactive mastery experiences. Enactive mastery experiences are the authentic tasks that, when completed successfully, contribute to building and strengthening one's sense of efficacy, and were determined to have positive correlations to achievement (Bandura, 1997; Dweck, 2006; Goddard, 2001). Goddard's (2001) study found "Mastery experience explained nearly two thirds of the variance between schools in collective efficacy... [whereas] SES and race were no longer statically significant predictors of differences among schools in collective

efficacy” (p. 474). Other research suggests that mastery experiences help build confidence and perceptions of skillfulness, which then effect efficacy levels and approaches to future learning (Bernacki et al., 2015; Usher & Pajares, 2006; Zimmerman, 2011). Bernacki et al. (2015) added nuance to the discussion: students modified their judgments on their abilities to solve mathematics problems over time as more information and feedback was provided. The suggestion that students use multiple inputs and variables of mastery experiences to inform their efficacy judgments even while engaging in learning indicates that observation of efficacy is complex. This indicates it has value beyond the scope of a pre and post-performance survey, as the minutiae of the learning process is not readily documented exclusively via this method.

The findings on mastery experiences have several implications for classroom instruction. The first is that cultivating student efficacy through mastery experiences has potential to mitigate predictive factors of SES and race on achievement. Second, teachers must also develop classrooms where agency is developed and employed by students in order to effectively build and cultivate efficacy. Lastly, students must be taught to set effective goals and how to filter and capitalize on quality feedback to foster a sense of accomplishment rather than discouragement. This concept is discussed later in the sections describing feedback, growth mindset and the regulation of emotions.

Vicarious experiences. Peer models and self-modeling affect efficacy beliefs in several different ways. Not only do peer models provide a standard to which students can assess their own capabilities, they can also teach skills to manage tasks, particularly when perceived inefficacy is present in observing students (Bandura, 1997). Bandura (1997) attributes this phenomenon to observers visualizing themselves as their peers in the model. If one views people similar to themselves succeeding at a task, the observer also feels they can succeed. Conversely,

if the observer sees a person failing at a task, their sense of efficacy can be lowered (Bandura, 1997; Boisvert & Rao, 2015). This is even more prevalent when the observer has no self-awareness of their own capabilities (Bandura, 1997).

The findings on modeling have several implications. The classroom implications for these findings include evidence that students need to understand their own capabilities through self-appraisal and reappraisal, and learn to self-diagnose strengths and areas for improvement. Boisvert and Rao (2015) determined self-modeling, watching videos of performances or deconstructing personal learning, accelerated the learning process in case studies of English Language Learners and contributed to their positive perceived self-efficacy. This also requires students know where they are, determine where to go, and make a plan for getting there in the form of goal setting.

Verbal persuasion. Feedback towards goals also has important implications in the classroom. Goal setting is an important part of the mastery process and is built on mastery experiences. Cognitive processing about one's capabilities relative to these goals occurs before, during, and after experiences, and students weigh and filter feedback from these experiences, choosing to integrate it into their sense of efficacy or discard it (Bandura, 1997). Goals requiring the development of requisite skills should not be judged on the amount of effort or hard work put towards its achievement; this can contribute to the demoralizing of students and undermine their sense of efficacy by replacing ability constructs with aptitude (Bandura, 1997). "In contrast, viewing ability as an acquirable skill fosters a resilient sense of efficacy, adoption of challenging personal aspirations, proficient analytical thinking, and performance attainments" (Bandura, 1997, p. 226). Rather, goals should be set to include subskills, and feedback on these should be

on the quality of the work and learning relative to mastery rather than quantity or pace of progress.

Whereas feedback on mastery experiences and past performances may influence efficacy levels, Goddard (2001) also acknowledged that the level of agency teachers felt that they had to make decisions and take action to solve problems also influenced their collective efficacy beliefs. Tasks that allow for student mastery and agency are an important component for feeding self and collective efficacy. Lodewyk and Winne's (2005) discussion of their study on task level efficacy and student achievement echoed Wiley and Voss's (1999) recommendations for teachers to create tasks that require true mastery of content through analysis of multiple sources of information that students have to synthesize to justify their thinking, rather than creating simple tasks that require them to regurgitate or memorize information. Phan's (2013) findings corroborated this. Students need to engage with "deep learning strategies," (p. 101) reflective thinking, feedback, and mastery goals, while determining and creating action plans for their own goals.

Physiological and affective states. The difficulties, setbacks, and failures experienced with these mastery experiences have also been shown to contribute to efficacy beliefs. Bandura (1997) explained that if students experienced repeated success and encounter struggle or failure, they could become discouraged (p. 80). However, unsuccessful experiences are not guaranteed to have an adverse effect. Rather, struggle allows people to learn how to handle their emotions related to unsuccessful experiences, and eventually view and manage them as opportunities (Bandura, 1997). It is a person's interpretation and integration of this feedback that determines the impact on self-efficacy (Bandura, 1997; Goddard, 2001).

Further Differentiating Efficacy

It is important to distinguish efficacy from similar theories and constructs about non-cognition to better understand what it is and is not. Efficacy is not the sense of control or power over life's events. It is not self-esteem, self-concept, or motivation. While it is considered non-cognitive or SEL, it is distinct from, yet influenced by, other constructs such as grit, growth mindset, and self-regulation. The following paragraphs explore these dichotomies further.

Locus of control. Rotter (1966) described one's belief that their actions affect the events in their lives as internal *locus of control*. Fatalistic or apathetic people tend to believe outcomes are predetermined and cannot be altered by their actions (Corcoran, Pettinicchino, & Young, 2011; Mirowsky & Ross, 2003). Unfortunately for many people in chronic poverty, a series of negative events in their lives might contribute to a sense of loss of personal control. As the sense of powerlessness develops and their locus of control declines it reinforces beliefs that a person has no ability to avoid pitfalls or effectively problem solve, leading to inaction and further fatalism or apathy (Bandura, 1997; Mirowsky & Ross, 2003; Rotter, 1966). While locus of control can be a barrier to action, it does not affect whether a person believes they can do something (self-efficacy), just whether or not they will do it.

Self-esteem. Self-esteem is largely concerned with one's overall self-worth. A student's self-worth could have very little to do with his/her ability to set goals and produce the desired outcomes (Mone, Baker, & Jeffries, 1995). For example, a 7th grade boy's self-worth could be driven by his status as a starter on the basketball team, how many subscribers he has to his YouTube channel, or whether or not his crush returns his feelings. These attributes are not predicated on his beliefs that he can set a goal and take action to meet it. Therefore, a person

could have low-self efficacy and high self-esteem or vice versa; there is no effect or relationship between the two.

Self-concept. Similarly, self-concept is a distinctly separate concept from self-efficacy (Bandura, 1997; Bong & Clark, 1999). Self-concept is the holistic view one holds of him or herself, and is usually dependent on comparisons of oneself to others (Bong & Clark, 1999). While it is also derived from and affected by a person's experiences, this act of comparison and self-appraisal has almost no predictive qualities for behavior. In other words, whereas people exhibit behaviors that correlate to their efficacy, self-concept holds no such correlation (Bong & Clark, 1999; Pajares & Miller, 1995).

Motivation. Bandura (1997) argued that motivation as described by White (1959, 1960) and Harter (1981) is also completely separate from self-efficacy. This concept of motivation based on the idea that a person's action is rooted self-satisfaction and pleasure for the purpose of continued satisfaction and pleasure (Harter, 1981; White, 1959, 1960). This has little connection to self-efficacy other than to say a student may derive some sort of pleasure from meeting goals or taking actions that result in intended outcomes. However, pleasure is not a requisite for the belief a person holds in their ability to attain desired outcomes (Bandura, 1997).

Although motivation itself is not efficacy, another description of motivation, mastery motivation, may contribute to overall self-efficacy. Yarrow, McQuiston, MacTurk, McCarthy, Klein, and Vietze (1983) developed the concept of mastery motivation based on the idea that motivation is observable and is expressed by students via their attentiveness, exploration, and perseverance toward tasks that result in competency in the area of study. While competency may be an end goal for some students, patterns of competency are developed through mastery motivations; mastery motivation and autonomy contributes to self-efficacy and self-efficacy

contributes to motivation (Bernacki, Nokes-Malach, & Alevan, 2015; Dweck, 2006; Lüftenegger, Schober, van der Schoot, Wagner, Finsterwald & Spiel, 2012; Pajares, 1996; Yarrow et al., 1983). The connections between mastery, autonomy, and efficacy are explored further in later sections.

Mindsets. Several studies have demonstrated growth mindset has a relationship with efficacy levels. Although conducted on adults, Karwowski (2013) found a creative mindset had strong correlation to the levels of creative self-efficacy and problem solving abilities. Recently, Rattan, Savani, Chugh, and Dweck (2015) created policy recommendations calling for the inclusion of academic mindset interventions in schools to help reduce achievement and motivation gaps in students. These recommendations were made based on the findings of several studies linking mindset to achievement.

Also promising, Paunesku et al.'s (2015) research established that simple mindset interventions in growth-mindset and sense-of-purpose improved students' course completion rates grades, overall grades, and thereby GPAs. These interventions involved only two 45-minute sessions and simply required students read articles on how the brain learns and grows from mistakes and apply that knowledge in a new scenario, and articulate how schoolwork helps them meet life goals. The intervention policy for schools, specifically, promoted by Rattan et al. (2015) include the explicit instruction on how learning happens and a growth mindset, and structures for creating inclusive environments, and fostering a sense of belonging in students. For schools and teachers, this also means adopting curriculum, grading policies, and instructional practices that promote process orientations, risk taking, and growth (Rattan et al., 2015).

As previously mentioned, California's CORE districts began a large scale study of the SEL of almost half a million students in grades 4-12, in an effort to understand the relationship

between SEL and achievement, SES, gender, age, race, and school. The specific SEL skills measured by the study included growth mindset, self-management, and social awareness, and self-efficacy. Preliminary findings by Krachman et al. (2016) indicated growth mindset improved as students got older, as relatively equal for males and females (females being slightly higher in all grades), and along with self-efficacy and self-management, was a strong predictor of student achievement in high school (see Figure 2).

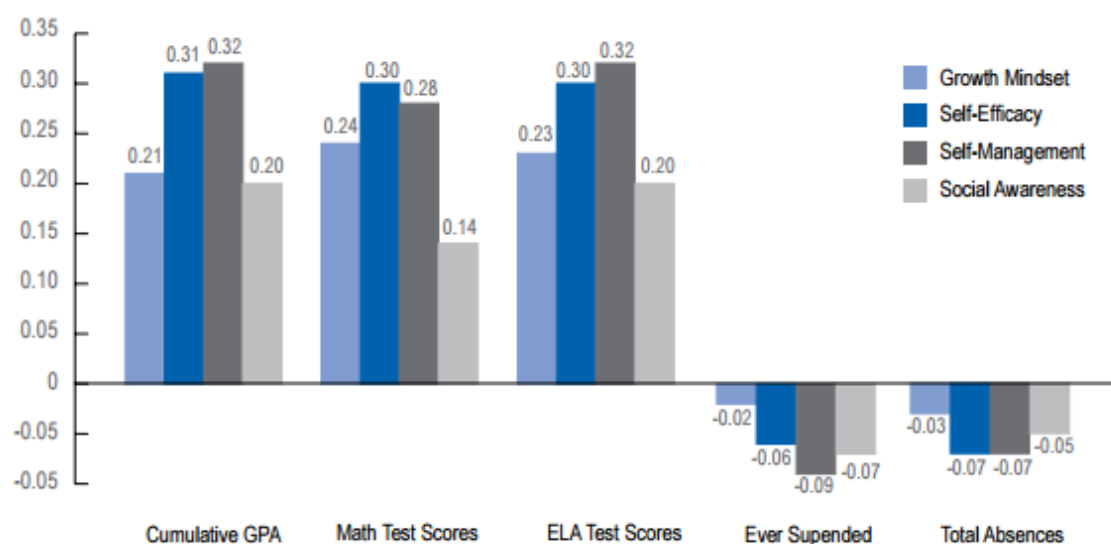


Figure 2. The CORE District SEL survey results for grades 4-12 as compared to academic and behavioral outcomes as it appears in Krachman, Arnold, and Larocca (2016, p. 15).

Autonomy, self-regulation and self management. Self-regulation is a self-directive process through which students are able to manage their thoughts and behaviors to transform mental abilities into academic skills (Pintrich & De Groot, 1990; Schunk & Zimmerman, 1998; Zimmerman, 2002;). Bandura et al. (2003) noted “the capacity for self-regulation is one of the core features of human agency in social cognitive theory” (p. 769) and earlier described the goal of education as the means to “equip students with the self-regulatory capabilities that enable them to educate themselves” (Bandura, 1997, p. 174). Zimmerman (2002) also noted its

importance in generating effective life-long learners. Self-regulation occurs because students are proactive learners and generate beliefs, thoughts and behaviors that result in the attainment of their goals. They are generally aware of their strengths and weaknesses, and actively work to improve, especially during informal or unstructured learning times, while continuing a high degree of self-management (Zimmerman, 2002, 2011).

Self-regulation is not an accident or a fixed trait. It can be taught through modeling and intentional instruction by teachers and peers. Students who self-regulate proactively seek out guidance and feedback to make adjustments and improve their learning (Bandura, 1989; Zimmerman, 2002, 2011; Goetz, Nett, & Hall, 2013). The self-regulative process students engage in when encountering a task are: setting goals, adopting strategies, monitoring their progress, restructuring their environment or context to be more conducive to learning, managing their time, self-evaluating, attributing causes to their effects, and making adaptations in the future (Zimmerman, 2002, 2011; Goetz, Nett, & Hall, 2013). The quality of self-regulation is to a certain extent determined by their level of self-efficacy, as self-efficacy influences each step of the process. Further, students' self-efficacy can impact future agency and self-efficacy. This cyclical relationship is shown in Figure 3 (Zimmerman, 2002, p. 67).

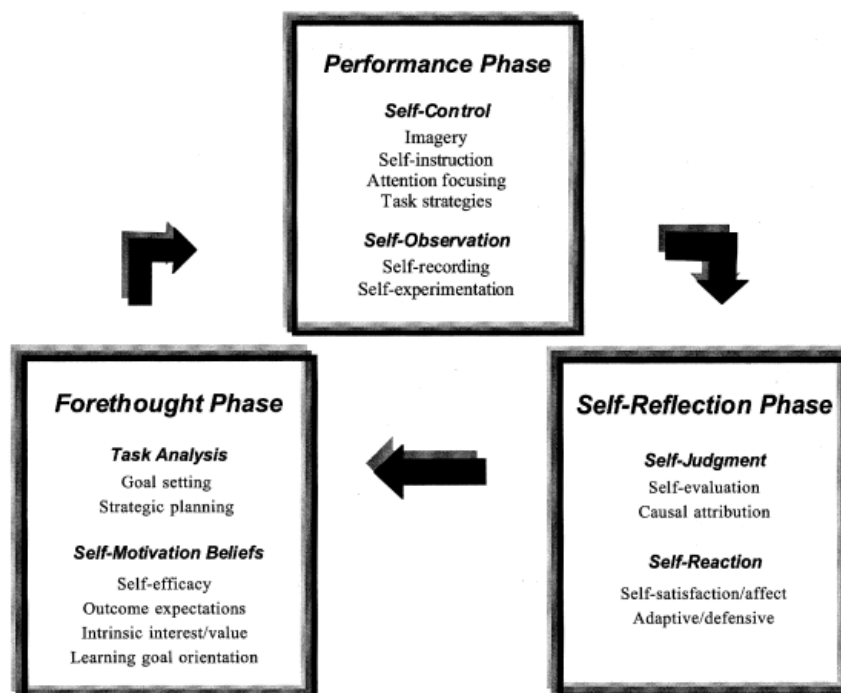


Figure 3. “The phases and subprocesses of self-regulation” as it appears in Zimmerman (2002, p. 67) from J. E. Davison and R. Sternberg (Eds) (2003).

Furthermore, several studies have suggested that the level of autonomy and the opportunities for self-regulation and management students are afforded has a positive relationship with their level of agency and efficacy (Lüftenegger et al., 2012; Putney & Broughton, 2011; Reeve, Ryan, Deci, & Jang, 2008; Zimmerman, 2002). Lüftenegger et al. (2012) surveyed 2,266 middle school students in Austria and determined their academic efficacy had a positive correlation with the level of perceived autonomy they had in class. The students’ perceptions on the instructional strategies utilizing self-reflection were shown to have a positive and significant relationship with students’ efficacy levels. Interestingly, Lüftenegger et al. (2012) found only moderate relationships between autonomy, efficacy, and academic grades. This could indicate that using grades to measure academic achievement may not be the best method. The researchers also recognized the limitations of their survey methods and

recommended future studies involve teacher interviews and classroom observations for validation of their findings. The use of observations and interviews to document self-regulation was also noted by Boekaerts and Corno (2005) as an effective method for data collection.

Grit. Duckworth (2012) defined the term ‘grit’ as “perseverance and passion for long-term goals” (p. 1087). But the concept of grit has come under fire recently due to the perception that it devalues the persistence and grit that minority students and those from disadvantaged backgrounds display regularly in their personal lives (Socol, 2014). Distinctions between academic grit and personal grit have begun to emerge in narratives, but research on the relationship to grit and student achievement has been unable to establish a firm correlation. In a recent study examining the relationship between academic resilience and self-efficacy, Cassidy (2015) ascertained that having high levels of self-efficacy resulted in increased academic resilience from students. This was established through vignettes of experiences that were admittedly limited in the traumatic effects they could have on subjects (Cassidy, 2015). It may be that the scenarios were not sufficient for inducing meaningful levels of stress that would activate or deactivate academic resilience. Similarly, in a meta-analysis of grit research, Credé, Tynan, and Harms (2016) found little evidence that grit is predictive of academic success. They too noted limitations to the current measures of grit and perseverance and its impact academic performance. For the purposes of this study, this would indicate that while resilient behaviors may indicate efficacy is present, their presence should not be the sole indicator of efficacy due to the low to moderate correlation grit has on achievement and the lower likelihood that they alone cause improvements in academic performance.

The combination of several other SEL indicators and their relationship with academic achievement was the subject of several recent studies. West et al., (2016) examined

conscientiousness, self-control, grit, and growth mindset and their relationship with academic achievement. Specifically, they were looking for differences in these relationships in charter schools in Boston as compared to their traditional school counterparts. West et al. (2016) surveyed more than 1,300 students in eighth grade using the Big Five Inventory (John & Srivastava, 1999) to measure conscientiousness, the Impulsivity Scale for Children (Tsukayama, Duckworth, & Kim, 2013) to measure self-control, the Short Grit Scale (Duckworth & Quinn, 2009) to measure persistence, and a three item set (Dweck, 2006) to measure growth mindset. Interestingly, West et al., (2016) learned that the students in charter schools, while generally making larger gains in state test scores than their district peers, had lower levels of conscientiousness, self-control, grit, and growth mindset, and scored lower in fluid reasoning skills. The researchers also found that all of the charter schools emphasized “no-excuses” as a means of character development which could contribute to negative impacts on student test scores and reasoning.

West et al. (2016) noted two corresponding issues in their methods that that could have contributed to misreporting and incongruent results. Reference bias in survey responses may have contributed to the differences in interpretation of the questions and in their expectations for themselves, which would influence how they respond to questions. Specifically, it could have been that students in the charter schools “use a higher bar when assessing their own conscientiousness, self-control, and grit when they attend schools that establish high expectations for student effort and a ‘no excuses’ disciplinary culture” (West et al., 2016, p. 163). Duckworth and Yeager (2015) also noted concerns about the limitations of surveys, and although Yeager et al. (2013) acknowledged variance in survey results due to self-comparisons to peers is still valid and predictive of academic outcomes within a school, it could be problematic when making

comparisons between schools. West et al. (2016) acknowledged that the academic and behavior expectations between schools made it difficult to compare one school to another and pushed for better measurement methods than the student self-reports.

MESH. As previously mentioned, California's CORE Districts have begun a large-scale study of the level of the Mindsets, Essential Skills, and Habits (MESH) of almost half a million students in grades 4-12, and their relationship to achievement, SES, gender, age, race, and school. The specific SEL skills measured by the study include growth mindset, self-management, and social awareness, and self-efficacy. Preliminary findings indicated growth mindset improved as students got older, and was relatively equal for males and females (females being slightly higher in all grades). Females also tended to have higher levels of self-management, and although levels dipped for both genders in grades 7-10, they returned to prior levels by eleventh and twelfth grades. Self-management was the best predictor of ELA achievement in elementary and middle school (Krachman et al, 2016).

Females tended to have higher levels of social-awareness as well. In this case, as students got older their levels declined starting in fourth grade, bottomed out in ninth grade a half of a point lower, and only improved a tenth of a point by twelfth grade. Females tended to have higher levels of self-efficacy until sixth grade when their scores went on a rapid decline of .5 between sixth and ninth grades while males declined by only .25 in the same period. Furthermore, students in poverty scored lower than their more affluent peers in every indicator, and African American scored lower in self-management and social awareness than their Caucasian peers (CORE Districts, 2016b).

Correlations Between Academic Efficacy and Student Achievement

Over the past few decades, research has also overwhelmingly demonstrated a correlation between self-efficacy and academic achievement (Farrington et al., 2012; Foster, 2015; Mega, Ronconi, & De Beni, 2014; Multon, Brown, & Lent, 1991; Schunk, 1989; West, 2014; Zahodne, Nowinski, Gershon, & Manly, 2015; Zimmerman, 1999). Many of the studies explored in the following paragraphs determined that academic self-efficacy erased the education or cognitive gap in students and adults in multiple situations.

There are several levels of academic efficacy. Table 1 illustrates the levels of specificity and their correlation to student achievement. It is important to note that multiple studies have found that the more participants were able to connect efficacy levels to specific domains, topics, activities, or tasks, the stronger the correlation to achievement (Bandura, 1997; Bong, 2001; Bong & Clark 1999; Bernacki, 2013; Foster, 2016; Lynch, 2013; Pajares, 1996, 2003). Studies involving task-level or task-value efficacy have demonstrated the strongest correlations of achievement to efficacy levels. Furthermore, there are distinctions between self and collective efficacy; each are explored in subsequent sections.

Table 1
Examples of the Levels of Specificity and Their Correlation to Achievement

Level of specification	Example	Correlational value
General	How one feels about their ability to successfully accomplish tasks as a person.	low
Setting	How one feels about their ability to successfully accomplish tasks related to academics, health, relationships, etc.	low
Domain	How one feels about their ability to successfully accomplish tasks in mathematics, writing, soccer, etc.	moderate
Situation	How one feels about their ability to successfully perform on a specific section of the ACT, in a specific game, etc.	high
Task	How one feels about their ability to successfully accomplish a task such as multiplying fractions, writing dialogue, dribbling a ball, etc.	high

Self-efficacy and achievement. Earlier studies tended to focus on the direct correlation between self-efficacy and achievement. In a study by Bouffard-Bouchard et al. (1991), students were given feedback that they either achieved at higher or lower rates than their peers, regardless of their actual performance on a given task. Those who received the feedback that they exceeded their peers reported higher levels of self-efficacy achieved at higher levels, set loftier goals, and used more efficient strategies than those who were told they underperformed regardless of actual cognitive ability or performance.

That same year, Multon, Brown and Lent (1991) conducted a meta-analysis of 40 self-efficacy studies established effect sizes of .38 for student performance and .34 for persistence suggesting that “self-efficacy beliefs account for approximately 14% of the variance in student’s academic performance and approximately 12% of the variance in their academic persistence” (p. 34). These early studies helped establish that a person’s level of self-efficacy contributes to his

or her behavior in the process of achievement, and can be fostered and grown (Bandura, 1997; Bouffard-Bouchard et al. 1991; Multon et al., 1991).

More recently, Zahodne, Nowinski, Gershon, and Manly (2015) conducted a study on over 1000 adults in which various executive functions including memory, vocabulary, processing speeds, and efficacy were assessed for their relationship to the participants' education levels. They discovered that participants with low levels of education and high self-efficacy achieved cognitive performance at levels similar to those with high levels of education. In other words, a person's belief that they can do something was a more powerful predictor of their success than whether or not they had the requisite knowledge, skills, or cognitive ability (Bouffard-Bouchard et al., 1991; Schunk, 1989; Zahodne et al., 2015).

In addition to increased cognitive achievement, other added benefits to high levels of academic self-efficacy include lower levels of depression and a smaller likelihood of dropping out of school which tend to plague low SES students more than their affluent peers (Quiroga, Janosz, Bisset, & Morin, 2013). Students with higher levels of academic efficacy have also been shown to have higher levels attendance, and lower rates of suspension (Quiroga et al., 2013; West et al., 2016). It is reasonable to assume based on this research a person in chronic poverty with high levels of self-efficacy can become an agent of change with the ability to improve his/her life and move its trajectory to a completely different path (Bandura, 1997). This is a powerful construct with formidable implications for teacher effectiveness in low SES settings as in examined in later sections.

Collective efficacy and achievement. Many research studies discovered positive correlations between collective teacher efficacy and student achievement (Goddard, 2001; Goddard, Goddard, Kim & Miller, 2015; Goddard, Hoy & Woolfolk Hoy, 2000; Hoy et al.,

2002). While conducting research on teachers' collective efficacy in urban schools, Goddard, Hoy, and Woolfolk Hoy (2000) found that collective efficacy was a better indicator of student achievement than SES. One suggested reason for this is that strong levels of collective efficacy are associated with teachers who have high expectations, are more persistent, and overcome failures more readily (Goddard et al., 2000). Furthermore, within the efficacy loop, these actions breed a culture where the perception of the collective group capability then creates pressure for all teachers to perform and the perception that the majority do so successfully. This then has generative impacts on students and their levels of efficacy and achievement (Bandura, 1997; Goddard et al., 2000). Although collective efficacy is typically used to describe the group-level efficacy of the teachers or leaders of a school, the concept of collective student efficacy may play a parallel role in the culture of a classroom.

Studies examining the relationship between students' collective efficacy and their achievements are rare and surveys are still in their infancy (Pina-Neves, Faria, & Rätty, 2013). Although Putney and Broughton (2011) did an extensive analysis on the collective classroom efficacy levels of students, the researchers did not examine the relationship between them and student achievement. Pina-Neves et al. (2013) on the other hand, conducted an analysis on the relationship between individual and collective efficacy levels of 385 high school students in Portugal, and their relationships to the students' achievement. The researchers' results were mixed. Individual efficacy levels of students had a stronger correlation to both mathematics and Portuguese, although a second study found collective classroom efficacy had a stronger relationship to student achievement in Portuguese (Pina-Neves et al., 2013). They also found males and students at public schools had stronger correlations between their collective efficacy levels and achievement.

Pina-Neves et al. (2013) noted the highly competitive nature of Portuguese schooling and attributed their findings to highly individualized achievement measures, with private school students more motivated by competitive achievement goals. However, a tremendous limitation of Pina-Neves et al.'s (2013) study was the scale used to measure students' collective efficacy, which asked students' questions regarding their efficacy at the setting level. The individual self-efficacy scale was domain specific (which, according to Table 1, has only moderate correlation value). The researchers might have anticipated the lower correlative value between the results of the setting-specific collective efficacy survey and achievement, as compared to the domain-specific individual collective efficacy results, which likely impacted their findings (Bandura, 1997; Bong, 2001; Bong & Clark 1999; Bernacki, 2013; Foster, 2016; Lynch, 2013; Pajares, 1996, 2003).

Correlations with the Learning Environment

How teachers structure their classrooms and learning environments (the contexts in which students learn) shapes the culture of the classroom, and according to Fraser & Walberg (1991) is one of the most important determinates of student learning. The correlation between the learning environment and student achievement has been verified by repeated studies (Connor, Spencer, Day, Giuliani, Ingebrand, McLean, Morrison, 2014; Davis & Warner, 2018; Dorman, 2001). Furthermore, studies have also found a correlation between the classroom learning environment and students' academic self-efficacy levels (Dorman, 2001; Yerdelen and Sungur, 2019). Dorman's (2001) study measured the correlation between the classroom environment, and students' academic efficacy levels. His research was conducted in Australia on 1055 students in Grades 8, 10, and 12. His methods included surveying students about their perceptions of their learning environment via seven scales from the What Is Happening in This

Classroom? (WIHIC) survey (Aldridge & Fraser, 2000) and three scales from the Constructivist Learning Environment Survey (CLES) (Taylor, Fraser & Fisher, 1997). Dorman (2001) determined that there was a statistically significant, positive correlation between students' level of academic efficacy and their learning environment ($p < .001$). Regression indicated the seven scales from the WIHIC survey accounted for 35% of the variance in student self-efficacy measures. The CLES was a less reliable predictor at 7%.

Yerdelen and Sungur (2019) confirmed these results in their cross-sectional study of 8198 seventh grade students and 372 of their science teachers in Turkey. Participating students were administered the WIHIC survey to measure their perceptions of the learning environment, as well as the Motivated Strategies for Learning Questionnaire and the Achievement Goal Questionnaire. Their research found the classroom learning environment was significantly correlated with students' self-regulation ($p < .001$) and self-efficacy levels ($p < .05$). Yerdelen and Sungur (2019) recognized their reliance on student and teacher perceptions via survey as a limitation of their study and recommended future researchers expand their collection methods to include observation to better understand the relationships between variables.

Correlations with Peer Relationships

Positive correlations between students' self-efficacy levels, academic achievement outcomes, social competencies, and peer relationships have been found in several studies (Ladd, 1990; Masten & Coatsworth, 1998; Patrick, Hicks, & Ryan, 1997; Schwartz, Gorman, Nakamoto, & Toblin, 2005; St Clair-Thompson, Bugler, Robinson, Clough, McGeowin, & Perry, 2015; Tulis, Und, & Dresel, 2018). Patrick et al. (1997) administered the PALS (Midgley et al., 1996) survey to measure academic self-efficacy, the Harter (1982) Perceived Social Competence Scale to measure students' perceived social efficacy, and the Responsibility Goal

Scale (Wentzel, 1991) to measure academic social responsibility to 753 fifth graders in 35 classrooms in Michigan. These data were compared to their final grades at the end of the semester. They found that peer and student to teacher relationships were positively correlated with students' academic outcomes and levels of self-efficacy. However, they noted the limitation of their study for better understanding the way teachers influence student self-efficacy levels and recommended further research in this area.

Implications for this Study

There are several potential indicators that teachers and observers should consider when determining whether an agentic and efficacious culture is present or has potential for cultivation. If teachers not only consider and receive feedback on these indicators in their classroom to monitor and adjust their instruction, thereby intentionally cultivating the requisite beliefs, the efficacy loop can be propagated (Bandura, 1997; Multon et al., 1991). Because agency and efficacy affect the way in which students think and behave, they have predictive qualities for students' cognitive and metacognitive behaviors (Pajares, 2008). Specifically, students with a high sense of agency have been determined to exhibit the following observable behaviors:

- Manage time effectively (Bouffard-Bouchard et al., 1991; Zimmerman 2002, 2011).
- Transfer knowledge and skills and select appropriate strategies to apply these in new situations (Bouffard-Bouchard et al., 1991; Zimmerman, 2002).
- Demonstrate perseverance and problem solving, including asking for help, when faced with barriers to success (Bandura et al., 2003; Bouffard-Bouchard et al., 1991; Ryan, Gheen, & Midgley, 1998).

- Are less likely to reject good solutions prematurely (Bouffard-Bouchard et al., 1991).
- Take academic risks by taking on more challenging tasks, setting higher standards and setting more challenging goals for themselves (Bandura, 2012; Pajares, 2008; Putney & Broughton, 2011).
- Achieve their goals (Pajares, 2008).
- Use more efficient strategies (Bandura, 2012).
- Achieve at higher levels intellectually (Bandura, 2012).
- Incorporate self-reflection, metacognition, and growth-focused language while owning their decisions and the consequences via strategic thinking (Krachman et al., 2016; Pajares, 2003; Pintrich & De Groot, 1990; Putney & Broughton, 2011).
- Use metacognition to monitor their understanding and the relevance of the content (Moos, 2014; Pintrich & De Groot, 1990; Putney & Broughton, 2011).
- Self-regulate during the process of learning by monitoring their own effort, work quality, and learning habits (Bandura, 2012; Pintrich & De Groot, 1990; Putney & Broughton, 2011; Zimmerman, 2011) .

I expected to find similar student behaviors in my observations of classrooms with higher levels of student agency and efficacy, and used this list when coding my qualitative data.

If evaluations systems are going to meet the goals of providing support and helping teachers prioritize and develop in areas where student achievement is most impacted, then a culture of agency and efficacy may need to be included in this process. The implication for teacher evaluation is that raters should observe tasks and conduct discussions with students in the process of learning in order to capture an accurate view of and context for the level of collective

efficacy in the classroom. An evaluation tool can and should remain learner centered and effectively capture evidence of observable agentic and efficacious student behaviors to measure effectiveness and provide feedback to effect student achievement.

Summary

In summary, students' task-level efficacy is a powerful predictor of their success than whether or not they have the knowledge or skills required to complete the task or goal (Bouffard-Bouchard et al., 1991; Schunk, 1989). Moreover, a classroom culture of efficacy and agency could potentially demonstrate a relationship to student achievement if the central tendencies used to indicate agency and efficacy are applied as in Goddard et al. (2000), and surveys, observations, and discussion/interview research methods are utilized (Putney & Broughton, 2011). Chapter 3 presents my concurrent triangulation mixed-method design for measuring the culture of agency and self-efficacy (CASE) in classrooms.

CHAPTER 3

RESEARCH METHODS

This chapter describes the concurrent triangulation mixed-methods design of this study. It begins with an explanation of my research design followed by details of the demographics of the sample and descriptions of the instruments. It progresses through an accounting of the data collection procedures in the quantitative and triangulation qualitative portions of the study. It concludes with a description of the data analysis process and the limitations of the study.

Purpose of the Study

The purpose of this study was to explore factors related to Social Cognitive Theory that influence achievement and student success using concurrent triangulation mixed methods (Creswell, 2014). This study identified the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in high poverty classrooms. It is my hope these findings could be used to better support teachers and evaluation systems in the future. In the first quantitative phase of the study, the research questions explored the relationship between student achievement, perceived student agency and academic efficacy in classrooms, and the students' levels of perceived teacher effectiveness. The independent variables in this phase were demographics, perceived self-efficacy, self-management, and social awareness, student perception of teacher, and teacher perception of student as measured by student and teacher surveys. The dependent variables were academic achievement in ELA and mathematics as measured by Acuity (McGraw Hill) assessments. The purpose of the concurrent, qualitative Phase II of the study was to identify characteristics of classrooms and practices of teachers that establish agentic behaviors and academic self-efficacy beliefs of students.

Research Questions

The research questions this study attempted to answer were:

Quantitative Phase:

1. What are the relationships among students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness?

Qualitative Phase:

2. What are the observable behaviors and norms of students in classrooms with varying levels of cultures of agency and self-efficacy (CASE)?
3. What relationship do students' agentic and efficacious classroom behaviors have with their perceived self-efficacy levels and academic achievement?

Overall:

4. What are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?

Research Design

This study was a concurrent triangulation mixed methods design (Creswell, 2014; Greene & Caracelli, 1997; Tashakkori, Teddlie, & Teddlie, 2003) integrating an instrument-building, sequential exploratory model within the qualitative phase (Creswell, 1999; Creswell, Plano Clark, Gutmann, & Hanson, 2003; Creswell & Plano Clark, 2017). Two simultaneous phases of data collection addressed two sets of research questions united by an integrated research question. The mixed-method nature of this design provided a deeper perspective on the subject than a single method would allow, and produced a richer description of the results (Tashakkori et

al., 2003). Furthermore, while the qualitative methods allowed me to develop an observation tool from induction through deduction, triangulation corroborated results with Phase I data and allowed me to establish reliability (Creswell & Plano Clark, 2017). Figure 4 is a visual representation of my study containing annotations in relation to the four criteria attributed to mixed methods research design: implementation, priority, integration, and theoretical perspective (Creswell et al., 2003; Morse, 2003). These are explained more fully in subsequent paragraphs.

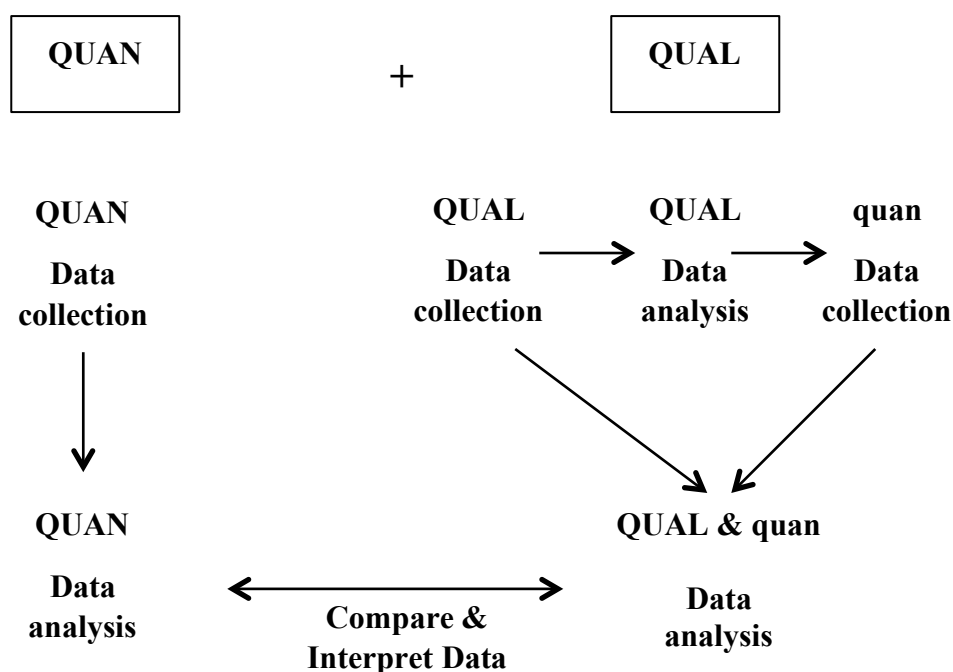


Figure 4. The concurrent triangulation design with an instrument-building, sequential exploratory model embedded in the qualitative phase.

Implementation refers to the timing of the data collection and relationship of the samples. Morse's (1991) notation system provided the field with a language for the timing of the data collection, namely simultaneous and sequential. Simultaneous collection denotes quantitative and qualitative methods collection phases occurring during the same span of time; whereas sequential collection occurs one before the other, usually with the first data set informing the

second collection. These methods were annotated with a plus sign (+) for simultaneous and an arrow (→) for sequential (Morse, 1991). Onwuegbuzie and Collins (2007) modified these terms slightly to be concurrent and sequential while Creswell and Plano Clark (2017) referred to these as convergent and sequential. Each set of researchers retained the general original meaning. These annotations are used to indicate the relationship between all phases in the process of the research design, including planning, collection, analysis, and interpretation. The researchers noted a concurrent design was appropriate for triangulation studies (Onwuegbuzie & Collins, 2007) and is the primary research design for my study. Sequential design utilized in Phase II of my study is typical for instrument-building models, beginning with an exploratory or inductive method, progressing through analysis, then instrument creation (Bloomberge & Volpe, 2016; Creswell, 1999; Creswell & Plano Clark, 2017).

The relationship of the samples are typically identified as identical, parallel, nested, or multilevel and are used to describe the correlation between the sample populations of the quantitative phase relative to the qualitative (Creswell & Plano Clark, 2017; Onwuegbuzie & Collins, 2007). As implied, an identical relationship between sample groups indicates the exact same participants comprise both. A parallel relationship means the participants in the qualitative and quantitative sample groups are different but from the same population. A nested relationship involves using a subset of the participants in one phase from the larger sample population in the other. In a multilevel relationship, at least two subsets of samples are used for different phases of the study. My study utilized an identical sample of students, as well as an identical sample of teachers in the quantitative and qualitative phases.

Priority refers to the method of collection, quantitative or qualitative, that may dominate the research. Morse (1991) also developed an annotation strategy to delineate emphasis on the

type of collection. The methods were shortened to qual and quan. All capital letters denoted more emphasis, while lowercase signified less. Although the overall structure of my concurrent triangulation study placed equal priority on each method (QUAL + QUAN), within the instrument-building, sequential exploratory model design of my qualitative phase qualitative was emphasized over quantitative and was therefore annotated as QUAL and quan.

Integration signifies the point at which the data collected from the two methods are combined. This can occur during the interpretation or analysis phases (Creswell & Plano Clark, 2017). Concurrent triangulation research methods allow data from the two phases to be combined during both the analysis and interpretation phases. Researchers noted that transforming data to allow for effective integration during the analysis phase could present a challenge (Creswell & Plano Clark, 2017; Creswell, et al, 2003; Morse, 2003). One method is to transform qualitative data into quantitative in order to make comparisons (Creswell & Plano Clark, 2017; Creswell et al., 2003; Tashakkori et al., 2003). This is the process my study used. After the instrument development process, I was able to collect quantitative data at the classroom level and tag these cases with aggregated Phase I data for triangulation of my results. It should be noted that multiple researchers in Tashakkori et al. (2003) noted challenges may arise in explaining discrepancies in the data collected by one method as compared to another.

The theoretical perspective is the theoretical or conceptual framework that supports the research and interpretation of the results of a study. According to Creswell et al. (2003), this perspective is comprised of the personal stances and experiences as well as the researchers' professional focus and lens from which they view their study. All studies may have a theoretical perspective or conceptual framework from which they implicitly or explicitly operate, with

transformational designs requiring one. My conceptual framework of CASE is explained more fully in Chapter Two.

Description of the Sample

This study included a nonprobability sample (Creswell, 2014; Onwuegbuzie & Collins, 2007) of 10 classrooms spanning grades three, four, and five in two public charter schools located in the urban core of Indianapolis. The schools had populations of students with over 95% receiving a free or reduced price lunch, a proxy for poverty. All individual students' responses to surveys were tagged using their email addresses for the purposes of tracking and analyzing demographic data. The selected schools were two public charter schools within the same network. One was a growing K-5 elementary, and the other was a combined K-12 school.

While pure heterogeneous grouping per SES was impossible for my study given the homogenously low-SES demographic of the student population, students were assigned to classrooms in a stratified manner to create a balance of between those who struggled and excelled in behavior and academic categories, those in Special Education (SPED) with Individualized Education Plans (IEPs), and those learning English (ELL) with Individualized Learning Plans (ILPs). The schools had no formal tracking systems in place but did identify students as having low, moderate, and high academic or behavior needs at the end of the prior school year for stratified assignment to classrooms the following year. All students with ILPs and IEPs were educated with the general education population at or greater than 90% of the time, and their assignments to classrooms were stratified as well. No students were tracked into a high ability program or self-contained classrooms. There was, however, a contingency of students who looped with their teacher from the prior year at both campuses (Cistone & Shneyderman, 2004).

Because of the stratified assignment of students to classrooms, classes began the year balanced in terms of number of students who need intense, moderate, and light academic supports, behavior supports, language supports, or learning supports, and by gender. That buffered students' prior achievement as considerable variable impacting the degree to which the teacher promotes self-efficacy in this study, tempering the effects of non-random assignment found by Steinberg and Garret (2016). Past anomalies in the stratification occurred when a student transferred to another school and was replaced with a student from the waitlist with a different need. Historically, the schools had approximately a 10% school-wide attrition rate.

All of the teachers at the two schools had varying degrees of training on SEL particularly in the areas of growth mindset and self-efficacy over the past two years. The schools' evaluation process was formative in nature. According to Shute (2008), the purpose of the formative feedback process is to modify the receiver's behaviors and thinking. Administrators, instructional coaches, and teacher leaders were trained to conduct walk-throughs and formal observations using a research-based (Danielson, 2013; Marzano et al., 2014; Strong, 2011) rubric created by a committee of school stakeholders. Administrators, coaches, lead teachers, and the teachers themselves collected evidence of teacher proficiency during these informal walk-throughs, instructional coaching meetings, and two formal observations each year. Teachers had a pre- and post- conference and conducted a self-evaluation for each formal observation. At the end of the year, the collective evidence was used to create a formal summative evaluation written by the administrator, coach, and teacher. Goals were created and set by the teacher for the following year at the summative conference.

Demographics

Table 2 contains the demographic percentages of the total 3rd through 5th grade student populations at both Garfield and Washington Academies. These students at both academies were divided into two classes of 25 to 27 students. The total student population enrolled in grades 3-5 at Garfield Academy, as compared to Washington, was fairly equal with some notable exceptions; the percentage of enrolled White and Black students were opposite of each other at each campus, and the percentage of ELLs at Washington Academy was larger than Garfield. The 10 participating teachers were all White females with a minimum of three years of teaching experience.

Table 2
Demographic Population Numbers [and Percent] of Garfield Academy and Washington Academy Enrolled in 3rd, 4th, and 5th grades

Students	Garfield Academy	Washington Academy
Total	161	165
Hispanic	74 [46]	73 [44]
White	48 [30]	25 [15]
Black	26 [16]	59 [36]
Multi-racial	13 [8]	8 [5]
Free/ Reduced Price Lunch	153 [95]	162 [98]
ELL Enrollment	40 [25]	54 [33]
SpEd Enrollment	18 [11]	17 [10]

Instruments

The instruments used in this study were a combination of validated surveys from the CORE District SEL / MESH data collection initiative and the questions from the PALS student survey that measure teacher actions. The schools provided me with the students' academic achievement scale scores on their formative benchmarks, Acuity A, B, and C, for ELA and mathematics. During the qualitative Phase II, I created an observation tool to help triangulate my

findings. The following paragraphs describe each instrument from the context of their validation and reliability measures, and their use in this study.

CORE District SEL/MESH Survey

The CORE Districts MESH survey was utilized by eight California school districts to measure the Mindsets, Essential Skills, and Habits (MESH) of over one million students. Between 2013 and 2016, students in the eight CORE Districts were administered the survey and the aggregate data were used to identify lower and higher performing schools for mentoring and support purposes. Student subgroups included ethnic, disability, socioeconomic, and English Learners, while 77% were low income, 27% were English learners, and 13% were students with disabilities (CORE Districts, 2013).

The survey demonstrated validity and reliability by Harvard University's Center for Education Policy Research over the course of the last three years (~ 0.6 using Cronbach's alpha), as well as shown to correlate to academic outcomes (CORE Districts, 2016b). The CORE District MESH survey was comprised of four to nine questions in four core competencies: growth mindset, self-efficacy, self-management, and social awareness. With a student sample size of 378,456, the self-efficacy competency had the highest level of reliability of .87 (see Table 3 for the reliability of each competency). Questions on growth mindset were shown to be unreliable when administered to students below fifth grade, while all others were determined reliable for the lower grades; the survey for my study therefore did not include the questions on growth mindset (Transforming Education, 2016). Transforming Education (2016) noted the competencies may be measured and scaled independently of each other and maintain the validity of the survey. They also noted anchoring vignettes did not improve performance or validity and were therefore absent from my study.

Table 3
Reliability of CORE District MESH Survey Competencies Using Cronbach's Alpha

Competency	Reliability
Growth Mindset	.70
Self-Efficacy	.87
Self-Management	.85
Self-Awareness	.81
Combined	.89

Over the course of the three years of implementation of the CORE District MESH survey, there was a stronger relationship between students' self-ratings within a school than between schools. While this would suggest some level of reference bias, this was not a significant concern for researchers. They determined that the "overall relationship between MESH skills and other student outcomes is stronger than the within-school relationship," balancing concerns about between-school bias (Transforming Education, 2016, p. 16). Nevertheless, Transforming Education (2016) recommended using data gathered from the CORE District MESH survey for formative purposes only until further improvements in the measures and administration protocols have been conducted to eliminate bias.

Questions from three of the four areas of competency from the CORE District MESH survey comprised my survey: self-management, self-efficacy, and social awareness. Students reported how frequently they used various self-management strategies over the past 30 days using a Likert scale (ratings from 1-5, *Almost Never, Once in a While, Sometimes, Often, Almost All the Time*). For self-efficacy, they reported on their level of confidence at completing certain achievements in ELA and mathematics separately. These achievements included variables such as getting good grades and meeting learning goals and used a Likert scale (ratings from 1-5, *Not At All Confident, A Little Confident, Somewhat Confident, Mostly Confident, Completely Confident*). Finally, students reported on their thoughts and actions when with other people over

the last 30 days using various Likert scales appropriate for each scenario. Students were directed that they could use their behavior tracking charts when answering this portion of the survey. The only modification I made to the survey questions was to replicate the self-efficacy questions so that they were specific to ELA and mathematics in order to correlate them to subject specific student outcomes. The full survey questions for this study are located in Appendix A.

I also used the “Teacher report on students’ MESH competencies” from the CORE District survey for teachers to rank individual students’ Self-Management and Social Awareness levels (see Appendix C for the survey). Transforming Education (2016) found teacher survey ratings of student skills to have a statistically significant relationship ($p = .001$) with their achievement. However, the correlation between elementary students’ self-perception of self-management scores and their teachers’ scores were not internally reliable as measured by Cronbach’s alpha ($\alpha = .40$) in Transforming Education’s (2016) initial sample of 166 schools spanning grades K-12. The researchers attributed the low correlation with the inclusion of third and fourth graders self-perception variable of growth-mindset as compared to middle and high school students (Transforming Education, 2016). This growth-mindset measure was not included in my study.

PALS Student Survey

The purpose of administering a portion of the PALS survey was to identify the perceived teacher actions that have high correlations to student agency and self-efficacy. The questions utilized from the PALS survey were those that referenced “my teacher” and were derived from the Teacher Mastery Goal and Academic Press subscales (see Appendix B for the survey questions). Researchers at the University of Michigan developed the PALS survey and later validated it using a confirmatory factor analysis (Midgley et al., 2000). The goodness of fit

index (GFI) measured at .96, indicating a strong fit between the survey model and the covariance matrix (Midgley et al., 2000).

Over the course of 10 years, the survey was administered to low-to-middle income students, approximately 55% minority, in nine school districts throughout three Midwestern states at elementary, middle, and secondary levels. Although the original survey was comprised of five different student scales, these can be administered separately (Midgley et al., 2000). All survey questions utilized a Likert scale for responses.

The sections I used in my student survey were the Teacher Mastery Goal and Academic Press questions. These were included to measure the students' perceptions of the goals and norms the teachers emphasize in their classrooms. Each had a separate alpha score (see Table 4 for the reliability as documented by Midgley et al., 2000). The Teacher Mastery subscale measured the "students' perceptions that their teacher emphasizes engaging in academic work in order to develop competence" (Midgley et al., 2000, p. 14), whereas the Academic Press subscale measured the "students' perception that their teacher presses them for understanding" (p. 21). As previously mentioned, these questions were posed using a five point Likert scale ranging from *Not At All True* to *Somewhat True* to *Very True*. See Appendix B for the specific survey questions for each subscale. From the students' perspective, the combined questions from both the CORE District and PALS surveys were one continuous survey.

Table 4
Reliability of PALS Teacher Mastery Goal and Academic Press Survey Subscales Using Cronbach's Alpha

Subscale	Reliability
Teacher Mastery Goal	.83
Academic Press	.79

Acuity Assessments

The series of three Acuity assessments, an assessment product of McGraw Hill, were structured to predict the academic achievement of students on the state summative assessment, ISTEP. They were progressive in nature with the percentage of grade-level standards increasing with each subsequent assessment. The first assessment, Acuity A, was administered in a single statewide window in October; standards from the prior grade level comprised 66.5% of the content for each subject with just 33.5% from the current year. Acuity B was administered approximately two months later in early December with approximately 33.5% of content from the prior year's standards and 66.5% from the current year. Acuity C was administered just prior to the first summative state-testing window in mid to late February, with 100% of on-grade-level standards comprising the assessment. The allocation of grade level standards per assessment became a condition in this study that I will discuss in subsequent chapters.

CASE Observation Tool

During the qualitative instrument-building, sequential exploratory model phase II of my study I developed an observation tool I used to collect data during my final round of observations (see Appendix F for this final version). I used grounded theory to develop the observation indicators and collection metrics, and went through two iterations via open, axial, and selective coding prior to the last draft (Glaser & Strauss, 1967). More information on this process can be found in the Data Analysis section of this chapter and in Chapter 4. While all drafts developed in the qualitative phase helped to document agentic and efficacious behaviors in the classroom, I determined for the sake of clarity it was necessary to connect each to one of Bandura's (1997) sources of self-efficacy on the final tool to assist with data collection, analysis, and triangulation. I used my literature review to help with this process.

Rather than simply collect the number of instances of observed phenomena during an observation, I created a single four-point “ownership” scale that allowed observers to document whether most students were exhibiting the behavior independent of the teacher (3) doing so with teacher support or prompting (2), whether the behavior was mostly the responsibility of the teacher (1) or if it was not observed (0). I titled this the Task Experience / Responsibility Scale and used it to collect quantitative data reported Chapter 4.

I utilized a partner observer² during third and final rounds of observations to provide feedback and establish inter-rater reliability for the tool. My observation partner and I convened to discuss, familiarize, and orient us to the final version of the tool using the example data for the indicator codes from prior rounds of observations (Appendix G contains sample “look for” criteria form we used when training). Then we independently watched the same six video clips of classroom lessons to collect data using the tool, and conducted four live classroom observations together to ascertain whether we could have inter-rater reliability. I collected her independent responses and compared them to mine and found us to be in 80% agreement or better on each indicator (a 0.80 on Kappa’s Benchmark Scale) and were within one scale point of each other when not in agreement. For example, where my partner scored the live training observation a 3 on the *Celebrating Accomplishments* indicator, I scored it a 2 based on my observation of a small group where the students required heavy prompted and support from the teacher with the behavior. Not all indicators were observed during the observations and were noted with a score of “0” for the indicator.

² My partner observer was an instructional coach in the secondary program at Garfield Academy who was also enrolled in a doctoral program and was in the dissertation proposal-drafting phase. She had no evaluative responsibilities for the teachers observed and had prior training in teacher observation and data collection both from the school, prior employers, and through her doctoral program.

Data Collection

As a concurrent triangulation mixed methods study, my data collection occurred in two simultaneous phases. Phase I focused on the quantitative collection while Phase II was qualitative and followed an instrument-building, sequential exploratory model design (Creswell, 1999; Creswell & Plano, 2017; Creswell et al., 2003). Figure 5 shows the timing, process and structure of my data collection. The paragraphs in the following sections detail the data collection process for each phase.

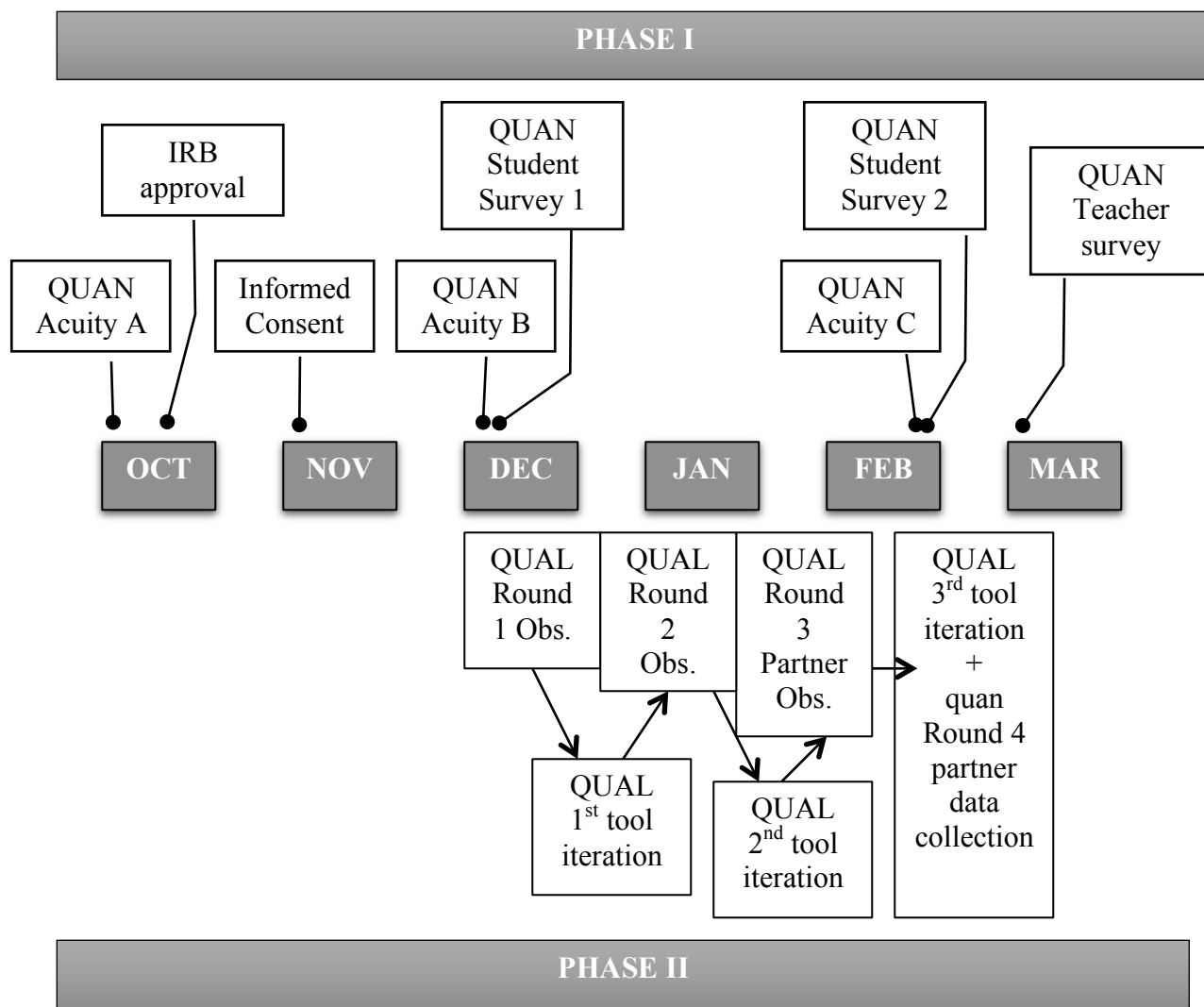


Figure 5. The timing, process and structure of Phase I (quantitative) and Phase II (qualitative) data collection. Acuity A data was administered by the teachers and provided to me upon the conclusion of Phase I as allowable in the data-share agreement.

Phase I – Quantitative Collection

During the quantitative phase of my study, I worked with school principals and teachers to administer the same cross-sectional survey twice, once in December and again in late February/early March, to 114 students in 10 total third, fourth, and fifth grade classrooms at two

schools. The participants were identified as students and teachers assigned to the same classroom at the same time for instructional purposes and were identical in each round of the collection. The rounds for survey administration corresponded with the administration of their formative assessment benchmarks, Acuity B and Acuity C. Due to the timing of IRB approval of this project, the survey was not administered after Acuity A (see figure 5). I timed this survey to capture the academic and agentic development of the students over the course of the year. I recognized that as the year progressed teachers had a greater amount of time to establish and reinforce the culture of their classrooms, and for students to know well what the teacher does and expects from them. The looped status of the participating classrooms supports this assumption as discussed in Chapter 4 and 5.

The purpose of the survey was to measure students' level of perceived self-efficacy, self-management, social awareness, and teacher effectiveness as an efficient method of data collection and for triangulation with Phase II observations. As discussed in the prior section, the survey contained questions from the self-efficacy, self-management, and social awareness competencies from the CORE District MESH survey, and the Perception of Teacher Mastery Goal and Academic Press subscale sections of the PALS survey. Students were emailed a custom, time sensitive link to the survey via Qualtrics at the start of each administration window, coupled with the their demographic and classroom data. Teachers assisted students with accessing and launching the survey³. To support the fidelity of the administration, videos containing the instructions for the survey were embedded in the survey. Read-aloud audio files for each question were also embedded to assist any student who may have needed reading

³ School officials provided me with student emails and demographic data in a spreadsheet for upload into Qualtrics. Only my dissertation chair and I had access to this data while it was stored in Qualtrics. I removed email addresses from the data file upon download prior to uploading it into SPSS. I deleted the survey and the protected data from Qualtrics to protect student privacy.

assistance. This helped reduce any teacher influence on student results as well as ensure continuity in the administration. The survey took approximately 15-20 minutes to complete.

At the conclusion of the study's collection window, I asked teachers to complete a survey on their individual participating students using the "Teacher report on students' MESH competencies" from the CORE District survey. They completed this just after the close of the Acuity C administration in early March. The survey (see Appendix C) specifically addressed individual students' Self-Management and Social Awareness. These results were later compared to the students' survey and achievement scores for triangulation purposes. More details on this analysis can be found in the analysis section of this chapter.

Phase II – Qualitative Collection

In the larger context of my study, the purpose of Phase II was to collect observation data to assist with and deepen the understanding of and explore the quantitative phenomena collected in Phase I as expected of an instrument-building, sequential exploratory model design, and to address the specific Phase II research questions (Creswell & Plano, 2017; Creswell et al., 2003). Rather than rely exclusively on surveys that better indicate students' perceptions of their behaviors in contrast to actual behaviors, researchers such as Bernacki, Nokes-Malach, and Aleven (2015) describe the importance of observing behaviors as they unfold in the learning process. Specifically, they recommended, "close observation and detailed recording of behavioral processes, as well as timely questioning about less overt processes" (Bernacki et al., 2015, p. 103). Furthermore, studies involving task-level or task-value efficacy have demonstrated the strongest correlations of achievement to efficacy levels (Bandura, 1997; Bong, 2001; Bong & Clark 1999; Bernacki, 2013; Foster, 2016; Lynch, 2013; Pajares, 1996, 2003). Although Bernacki et al., recommended observing participants individually, the context of my

study was classroom culture and evidence of agentic and efficacious student behaviors, and required observation of the collection of students as well as individuals. Therefore, Phase II of my study involved observing classrooms and informally conversing with individual students to better determine task level beliefs and behaviors within context of the collection of students. The following paragraphs provide more details about this process.

I observed classrooms using an inductive to deductive observational process comprised of a sequential cycle of observations with the purpose of exploring students' agentic and efficacious behaviors (Bandura, 1997; Tashakkori et al., 2003). The first collection window was an inductive process (Bloomberge & Volpe, 2016; Creswell, 1999; Creswell & Plano Clark, 2017). I observed each classroom once in either ELA or mathematics. While in classrooms I videoed activities, took notes on student and teacher behaviors and cultural features such as routines and norms, and scripted phrases from students and teachers. I also conducted brief and informal discussions with students, and noted items such as student work, handouts, and items on the walls (Creswell & Plano Clark, 2017). Pertinent information, such as the time of day of the observations, subject students were working on, pedagogical activities and structures, etc. were noted for analysis, and observations were focused mostly on documenting the activities and behaviors of students in the classroom that could be evidence of task-level agency and self efficacy (Bandura, 1997; Creswell, 2005; Merriam, 2009). Unlike typical grounded theory approaches, I did not transcribe the entirety of informal student discussions or videos, but used them to triangulate observations in my notes and later transcribed only specific illustrative phenomena for use in vignettes in Chapter 4 (Bloomberge & Volpe, 2016; Erickson, 1986; Moustakas, 1994; Saldaña, 2013).

In total, I conducted four rounds of observation. The first three were for the expressed purpose of inductive qualitative data collection and tool development. Each round resulted in an iteration of an observation tool that I used to collect data in the proceeding observation cycle. It was after the third round of observations that I developed a final iteration of the tool. After establishing inter-rater reliability on the final iteration of the tool prior to the final round to collect quantitative and qualitative data, my partner observer and I carried out a minimum of two 10-15 minute observations in each classroom. Time and not saturation, as defined by Strauss and Corbin (1998), determined how many observations were conducted and the duration of each. I observed each classroom in both ELA and mathematics subjects but my partner observed only one subject varied by classroom. The quantitative data were documented on the form for later analysis for correlations to student self-efficacy scores in each subject area as described in the following Data Analysis section of this chapter.

Data Analysis

Typical of concurrent triangulation design, my data analysis occurred separately for each phase and was integrated at interpretation. However, within Phase II, as is expected for an instrument-building, sequential exploratory model design, I conducted analysis after each round of classroom observations (Creswell & Plano, 2017; Creswell et al., 2003). Therefore, the analysis process of each phase of my study is described separately in subsequent sections.

Phase I - Quantitative Analysis

During the quantitative phase, student achievement data as well as student survey data were collected in two windows but not analyzed until the conclusion of the collection phase of the study. These data were integrated with the additional data from the teacher survey and the

linked demographic data of the students during the analysis phase. I used IBM's SPSS to run all statistical tests for these analyses.

I was provided participating students' Acuity achievement data from the school, which was reported in the form of students' individual scale scores for each subject for each assessment window. Scale scores are reported on a continuum where third graders are expected to score lower than fifth graders. Therefore, I had to calculate the z score ($z = \frac{(x - \bar{x})}{s}$) for inferential analysis based on grade level, rather than the combined population, otherwise my data would have been skewed with students in third grade statistically generally scoring below the mean and those in fifth scoring above. Therefore, I organized the data through split file according to grade and then ran the descriptive analysis for z scores. This allowed me to measure how many standard deviations below or above the local sample's grade level population mean a student's raw score was, as opposed to the total population. School-level data and not norm-referenced were used for this measure.

I used descriptive statistical tests to express demographic data listed in Table 2 and an inferential statistical t-test to conduct the within and between group analyses. T-tests, including Welch's test to confirm equality of variances, were conducted for the demographic independent variables against the academic achievement dependent variable z score. I also conducted an ANOVA test to determine whether significant differences existed between ethnic groups.

I used scores for the survey responses for individual question for each window, and calculated the means of the questions for each core competency (self-management, self-efficacy in ELA, self-efficacy in mathematics, perception of teacher, social awareness, and teacher perception of student) for each window. I used these as the independent variables and achievement z scores as the dependent variables in a series of student-level bivariate

correlational tests. This allowed me to detect any statistically significant relationships between students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness within and between the December and February administration windows.

Phase II - Qualitative Analysis

The instrument-design qualitative analysis process during Phase II of my study occurred sequentially, from inductive to deductive, to better understand and collect evidence of student and teacher behaviors related to classroom and task-level agentic phenomena (Bloomberge & Volpe, 2016; Creswell, 1999; Creswell & Plano Clark, 2017). I watched the videos of classrooms to add to my field notes and took additional notes of data added by pictures of classrooms and student work. As discussed previously, some portions of videos and informal discussions with students were later manually transcribed and pictures cataloged for additional evidence of the results of the study. I did not use a qualitative data analytical software package for this process.

As previously explained, I used a process in line with Grounded Theory (Glaser & Strauss, 1967) to conduct manual open, axial, and selective coding of my notes over the course of three rounds of observations. This process supported the drafting of the observation instrument. I used Holistic Coding to initially analyze the data and Provisional Coding using my literature review for the language of the codes and guidance on appropriateness and relevancy of the data (Bloomberge & Volpe, 2016; Saldaña, 2013). The codes (13) from the first round of observations were combined with Bandura's (1997) four sources of self-efficacy and four agentic properties to become the indicators on the first observation tool (see Appendix D for Tool 1). I then used a combination of Axial and Eclectic Coding (Saldaña, 2013) on my first and second

round of notes. I used these codes (18) to develop the second draft (see Appendix E for Tool 2). Finally, I selectively coded the data for Bandura's (1997) four sources of self-efficacy then reorganized and revised the indicators for the final draft (see Appendix F for Tool 3). I used the final draft during the fourth round of observations with my observation partner to collect quantitative data for triangulation with Phase I data (Creswell, 1999; Creswell & Plano Clark, 2017). This process and the results are explained in more detail in Chapter 4.

Using the quantitative data collected during the fourth round of observations, I tagged each classroom with their mean student data and the observation scores for classroom behavior indicators from the observation tool. I calculated the means of subscales (mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states) based in the individual observation indicators. I ran the bivariate correlation analysis between these data and the mean classroom-level data collected in Phase I to determine whether there were any significant correlations between them, and to triangulate my qualitative and quantitative findings. These findings are presented in Chapter 4.

Limitations

There are several limitations to this study. First, the sample size is small. In some cases, classrooms only had five participants. Second, the case study schools have little demographic variance between each other. They are very homogeneous in the SES of their student populations. While this study may be beneficial for schools with significantly high populations of low SES students, it would need replication in schools with more diverse and larger sample groups. Third, there was no prior relationship established between the CORE District survey and the PALS survey. This would need further study to determine if my results were typical. Fourth, the Acuity assessment used to measure student achievement is a predictive, progressive

assessment covering a diminishing percentage of standards from the prior year. I did not anticipate the impact this could have on my results. Fifth, the timing of the IRB process meant my study was truncated from the original timeline. The Phase II process was particularly short. Finally, the tool created during the last stage of Phase II will need to be revised and tested for further validity and reliability across a larger and more diverse sample. The analysis I conducted to help answer my third research question was on a sample size of 10 classrooms. Hox (2002) recommends at least 40-50 classrooms to effectively and reliably estimate group effects.

Summary

In summary, I conducted my concurrent triangulation mixed methods study in two phases. Phase I involved surveying 114 students and 10 teachers in 10 classrooms in grades three through five in two high poverty, urban schools. The quantitative survey results from the combined CORE District MESH and PALS surveys were used to measure the relationship between student self-efficacy levels and their perceptions of their teachers' effectiveness, and the teachers' perceptions of their students as further explored in Chapter 4. Phase II was an instrument-design, sequential exploratory model involving observations, informal discussions with students, and document reviews. The purpose of this phase was to capture and refine lists of agentic and self-efficacious student behaviors to develop iterations of an observation tool for future revision and larger-scale validation.

Chapter 4 contains the findings of the research conducted in both Phase I and Phase II. Specifically, the results are presented separately for each phase. Chapter 5 contains discussion of the comparative analysis of the findings of Phase I relative to Phase II, including my conclusions about the implications of this study and discussions of my overarching research question, "What

are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?”

CHAPTER 4

RESULTS

This chapter contains the results of my concurrent triangulation study and analysis of the data as these relate to my research questions on student agentic and self-efficacious behaviors and beliefs as an indicator of teacher effectiveness. This chapter begins with an explanation of the purpose and presentation of the research questions central to my study. Then I describe the sample and their demographics, including an analysis of the achievement data disaggregated by demographics. The remainder of the chapter is organized according to research questions. I begin with the Phase I quantitative research question one using the individual-level analysis of the student and teacher survey data and the students' academic achievement data collected in two windows. I then address the Phase II qualitative question two using the code notes and vignettes from my observations (Erickson, 1986). Finally, I present the results for my third research question by integrating and triangulating Phase I and Phase II data. The conclusions for my overall research question four are discussed in Chapter 5.

Purpose of the Study

The purpose of this concurrent triangulation mixed methods study (Creswell, 2014) was to identify the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in high poverty classrooms. In the first quantitative phase of the study, the research questions explored the relationship between student achievement, perceived student agency and academic efficacy in classrooms, and the students' levels of perceived teacher effectiveness. The independent variables in this phase were demographics, perceived self-efficacy, self-management, and social awareness, student perception of teacher, and teacher perception of student as measured by student and teacher surveys. The dependent variables were

academic achievement in ELA and mathematics as measured by Acuity (McGraw Hill) assessments. The purpose of the concurrent, qualitative Phase II of the study was to identify characteristics of classrooms and practices of teachers that establish agentic behaviors and academic self-efficacy beliefs of students.

Research Questions

The research questions this study responded to were:

Quantitative Phase I:

1. What are the relationships among students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness?

Qualitative Phase II:

2. What are the observable behaviors and norms of students in classrooms with varying levels of cultures of agency and self-efficacy (CASE)?
3. What relationship do students' agentic and efficacious classroom behaviors have with their perceived self-efficacy levels and academic achievement?

Overall:

4. What are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?

Description of the Sample

This study included a nonprobability sample (Creswell, 2014; Onwuegbuzie & Collins, 2007) of 10 classrooms spanning grades three, four, and five in two public charter schools located in the urban core of Indianapolis. The schools had populations of students with

over 95% receiving a free or reduced price lunch. All individual students' responses to surveys were tagged using their email addresses for the purposes of tracking and analyzing demographic data⁴. The selected schools were two public charter schools within the same network. One was a growing K-5 elementary, and the other was a combined K-12 school.

While pure heterogeneous grouping per SES was impossible for my study given the homogenous, low-SES student population; students were assigned to classrooms in a stratified manner to create a balance of between those who struggled and excelled in behavior and academic categories, those in Special Education (SPED) with Individualized Education Plans (IEPs), and those learning English (ELL) with Individualized Learning Plans (ILPs). The schools had no formal tracking systems in place but did identify students as having low, moderate, and high academic or behavior needs at the end of the prior school year for stratified assignment to classrooms the following year. All students with ILPs and IEPs were educated with the general education population at or greater than 90% of the time, and their assignments to classrooms were stratified as well. No students were tracked into a high ability program or self-contained classrooms. There was, however, a contingency of students who looped with their teacher from the prior year at both campuses.

Because of the stratified assignment of students to classrooms, classes began the year balanced in terms of number of students who need intense, moderate, and light academic supports, behavior supports, language supports, or learning supports, and by gender. That buffered students' prior achievement as considerable variable impacting the degree to which the teacher promotes self-efficacy in this study, tempering the effects of non-random assignment

⁴ School officials provided me with student emails and demographic data in a spreadsheet for upload into Qualtrics. Only my dissertation chair and I had access to this data while it was stored in Qualtrics. I removed email addresses from the data file upon download prior to uploading it into SPSS. I deleted the survey and the protected data from Qualtrics to protect student privacy.

found by Steinberg and Garret (2016). Past anomalies in the stratification occurred when a student transferred to another school and was replaced with a student from the waitlist with a different need. Historically, the schools had approximately a 10% school-wide attrition rate.

All of the teachers at the two schools had varying degrees of training on SEL particularly in the areas of growth mindset and self-efficacy over the past two years. The schools' evaluation process was formative in nature. According to Shute (2008), the purpose of the formative feedback process is to modify the receiver's behaviors and thinking. Administrators, instructional coaches, and teacher leaders were trained to conduct walk-throughs and formal observations using a research-based (Danielson, 2013; Marzano et al., 2014; Strong, 2011) rubric created by a committee of school stakeholders. Administrators, coaches, lead teachers, and the teachers themselves collected evidence of teacher proficiency during these informal walk-throughs, instructional coaching meetings, and two formal observations each year. Teachers had a pre- and post- conference and conducted a self-evaluation for each formal observation. At the end of the year, the collective evidence was used to create a formal summative evaluation written by the administrator, coach, and teacher. Goals were created and set by the teacher for the following year at the summative conference.

Participant Demographics

Table 5 contains the demographic percentages of the total 3rd through 5th grade populations relative to the sample populations for both Garfield and Washington Academies. The third through fifth students at both academies were divided into two classes of 25 to 27 students. The total student population enrolled in grades 3-5 at Garfield Academy as compared to Washington was fairly equal with some notable exceptions; the percentage of enrolled White and Black students were opposite of each other at each campus, and the percentage of ELLs at

Washington Academy was larger than Garfield. It should be noted that the percent of students who looped with their teacher in the sample group was more than 50% higher at Garfield Academy than Washington (see Table 5).

Table 5
Demographic Population Numbers [and Percent] of Garfield Academy and Washington Academy in 3rd, 4th, and 5th grades, and the Sample Groups

	Garfield Academy		Washington Academy		Combined
	3 rd – 5 th Grade Total (n = 161)	Sample (n = 62)	3 rd – 5 th Grade Total (n = 165)	Sample (n = 52)	
Students					
Hispanic	74 [46]	35 [57]	73 [44]	24 [46]	59 [52]
White	48 [30]	16 [26]	25 [15]	4 [8]	20 [18]
Black	26 [16]	6 [10]	59 [36]	21 [40]	27 [24]
Multi-racial	13 [8]	5 [8]	8 [5]	3 [6]	8 [7]
FRL	153 [95]	58 [94]	162 [98]	52 [100]	110 [97]
ELL	40 [25]	25 [40]	54 [33]	11 [21]	36 [32]
SpEd	18 [11]	9 [15]	17 [10]	5 [10]	14 [12]
Looped		37 [60]		5 [10]	42 [37]
Female		34 [55]		32 [62]	66 [58]

Note. Gender and looped data for the total 3rd - 5th grade population was not provided.

The sample from Garfield Academy included 62 students from five of the six available classrooms. The sixth classroom was eliminated from the study after the teacher unexpectedly went on early maternity leave. The Hispanic population of the sample group was larger than the cohort, while the White and Black populations were slightly lower. The percentage of ELL students was also higher in the sample than the cohort. The five participating teachers were in their third to eighth year of teaching and all were White females.

The sample student population from Washington Academy included 52 students from five of the six available classrooms. One teacher did not agree to participate in the study. The percentage of White students in the sample group was lower than in the cohort, while the percentage of Black students was higher. Unlike Garfield Academy, Washington Academy

experienced a lower participation rate for ELL students in their sample group than in the cohort. The five participating teachers were in their third to seventh year of teaching and all were White females.

Classrooms from all campuses were combined and then randomly assigned numbers to create anonymity and to protect against unintended bias in my analysis and interpretation phases. Table 6 contains the demographic breakdown of the sample group by classroom. Note that several classrooms had a very small number of participants ($n = 5$). This caused some of the demographic data to be skewed in classrooms with a smaller n size. The total N for the study was 114 students and 10 teachers. While small, it falls within the suggested 64 participants for one-tailed hypotheses and 82 participants for two-tailed hypotheses in correlational studies (Onwuegbuzie, Jiao & Bostick, 2004).

Table 6

Demographic Percentages of Participants by Classroom at Garfield and Washington Academies

	Garfield Academy					Washington Academy				
	Room 2 ($n = 6$)	Room 3 ($n = 8$)	Room 4 ($n = 17$)	Room 6 ($n = 16$)	Room 7 ($n = 15$)	Room 1 ($n = 7$)	Room 5 ($n = 5$)	Room 8 ($n = 17$)	Room 9 ($n = 5$)	Room 10 ($n = 18$)
Students										
Females	66.7	50	58.8	56.3	46.7	85.7	40	70.6	60	50
Non-White	50	62.5	76.5	81.3	80	85.7	100	82.4	100	100
FRL	66.7	100	94.1	100	93.3	100	100	100	100	100
ELL	0	37.5	52.9	50	33.3	14.3	0	35.3	20	16.7
SPED	16.7	12.5	17.6	0	26.7	0	0	11.8	20	11.1
Looped	100	0	0	100	100	0	100	0	0	0

Results

The remainder of this chapter is dedicated to the analysis of the data collected during my study. In review, this section begins with an analysis of the achievement data disaggregated by demographics followed by an analysis of the survey results by demographics. The remainder of

the chapter is organized according to research questions. I begin with the Phase I quantitative research question one using the student and teacher survey data and the students' academic achievement data collected in two windows for analysis. I then address the Phase II qualitative question two using the code notes and vignettes from my observations (Erickson, 1986). Finally, I present the results for my third research question by integrating and triangulating Phase I and Phase II data at the classroom level. The conclusions for my overall research question four are discussed in Chapter 5.

Demographics and Achievement

Academic achievement outcomes based on demographics was not explicitly addressed in my research questions. However, the correlation between student demographics and achievement is important to address prior to exploring the findings regarding my research questions. Therefore, the following paragraphs present the findings on student achievement relative to the student demographics presented in Table 6.

As mentioned in Chapter 3, Acuity scores are reported on a continuum with third graders expected to score lower than fifth graders. Because the achievement scores provided for this study were scaled scores, these had to be converted into contextually appropriate descriptive statistics. This required calculations of the z scores ($z = \frac{(x - \bar{x})}{s}$) for later inferential analysis based on grade level, rather than the combined school population. Otherwise the data would be skewed with students in third grade statistically scoring generally below the mean and those in fifth scoring above. Therefore, I split the data file by grade and then ran the descriptive statistics to accurately calculate z scores for student achievement on each Acuity administration relative to

the grade the student was assigned⁵. I was then able to conduct independent samples t-tests to compare Acuity achievement on the three ELA and mathematics tests, Acuity A, B, and C⁶, with the demographic variables.

For ethnicity, there was no statistically significant difference in the assessment scores of students by subgroup. The number of paid lunch students was extremely small at only four participants and was not a statistically significant predictor of achievement ($p = .846$). There were also no statistically significant differences in achievement between English Language Learners and native speakers (results including non-significant findings, can be found in Tables 18, 19, and 20 in Appendix H). Only gender, looped status, special education enrollment, and the enrolled school resulted in statistically significant differences between groups ($p < .05$). These results are presented in the following paragraphs.

I used independent samples t-tests to compare Acuity achievement categorical descriptive data on the three ELA and mathematics tests (Acuity A, B, and C) by the discrete demographic indicators. For gender, the results showed females ($M = -.06$, $SD = .842$) consistently scored lower than males ($M = .08$, $SD = .732$) on each assessment. However, this difference in achievement was not statistically significant with the exception of Acuity B in mathematics where *Welch's F* test confirmed a statistically significant difference between achievement on mathematics Acuity B for females and males ($F(1, 110.792) = 4.849$, $p = .03$).

Of the sample student group, 42 (36.8%) looped with their teacher from the prior year and 37 of these students attended Garfield Academy, comprising almost 60% of the total sample

⁵ Z scores were calculated relative to the school rather than a norm-referenced population.

⁶ Approximately 33% of the standards assessed on Acuity A were aligned to the grade level. The remaining 66% were from standards a year or more below grade level. For Acuity B, 66% of the standards assessed were aligned to the grade level. The remaining 33% were from standards a year or more below grade level. In contrast, 100% of standards for Acuity C were from the grade level.

population. There was statistically significant difference ($p < .05$) between students who did and did not loop with their teacher on every assessment. As presented in Table 7, analysis confirmed those who did not loop with their teacher performed lower than those who looped on every assessment in every subject.

Table 7
Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA and Mathematics by Looped Status

Group ^a	ELA			Mathematics		
	<i>M (SD)</i>	95% CI		<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>		<i>LL</i>	<i>UL</i>
Acuity A						
Looped ^b	.26 (.98)	-.04	.56	.36 (1.01)	.05	.68
Did Not Loop ^c	-.15 (.97)	-.38	.08	-.21 (.93)	-.43	.01
Acuity B						
Looped ^b	.32 (1.02)	.00	.64	.30 (.89)	.02	.58
Did Not Loop ^c	-.19 (.93)	-.41	.03	-.17 (1.01)	-.41	.06
Acuity C						
Looped ^b	.26 (1.02)	-.06	.58	.38 (.92)	.10	.67
Did Not Loop ^c	-.15 (.95)	-.37	.07	-.22 (.97)	-.45	.01

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit. ^a $N = 114$. ^b $n = 42$. ^c $n = 72$.

While there were a small number of students in the sample group enrolled in special education ($n = 14$), according to *Welch's F* test, there was a statistically significant difference in their achievement scores as compared to their non-special education peers on Acuity A ($p = .003$), B ($p = .001$), and C ($p = .001$) ELA assessments. Special education students performed lower than their peers on all three ELA assessments as presented in Table 8.

Table 8
Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA for Students According to Special Education Enrollment

Group ^a	<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>
Acuity A			
SPED ^b	-.57 (.64)	-.93	-.20
Non-SPED ^c	.08 (1.01)	-.12	.28
Acuity B			
SPED ^b	-.63 (.62)	-.99	-.27
Non-SPED ^c	.09 (1.00)	-.11	.29
Acuity C			
SPED ^b	-.50 (.45)	-.76	-.25
Non-SPED ^c	.07 (1.03)	-.13	.27

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit.

^a*N* = 114. ^b*n* = 14. ^c*n* = 100.

Analysis of school enrollment also indicated a statistically significant relationship with achievement on all three tests for both ELA ($p < .05$) and mathematics ($p < .001$). Students at Garfield Academy consistently outperformed their peers at Washington Academy on all Acuity assessments as presented in Table 9. It should be noted that this less than 10% of Washington Academy's sample population looped with their teacher compared to almost 60% at Garfield. Given the statistical significance of looping, this could be a contributing factor to the discrepancy between schools.

Table 9
Means, Standard Deviations, and Confidence Intervals of Acuity Outcomes in ELA and Mathematics Based on School Enrollment

Group ^a	ELA			Mathematics		
	<i>M (SD)</i>	95% CI		<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>		<i>LL</i>	<i>UL</i>
Acuity A						
Garfield ^b	.23 (1.01)	-.03	.49	.33 (1.06)	.06	.60
Washington ^c	-.28 (.90)	-.53	-.03	-.40 (.74)	-.60	-.19
Acuity B						
Garfield ^b	.18 (1.03)	-.08	.44	.35 (.85)	.14	.57
Washington ^c	-.21 (.91)	-.47	.04	-.42 (.99)	-.70	-.14
Acuity C						
Garfield ^b	.18 (1.06)	-.09	.45	.31 (.95)	.07	.55
Washington ^c	-.21 (.86)	-.45	.03	-.37 (.91)	-.63	-.12

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit.

^a*N* = 114. ^b*n* = 62. ^c*n* = 52.

Demographics and Survey Results

Survey results based on demographics was also not explicitly addressed in my research questions. However, correlations between student demographics and the survey results are important to address prior to exploring the findings regarding my research questions. Therefore, the following paragraphs present the findings on the two survey administrations relative to the student demographics presented in Table 6.

In preparation for this analysis, I calculated the means for each survey category on the CORE Districts survey (2016) and PALS survey (Midgley et al., 2000) based on the students' responses to the individual aligned survey questions for each administration. The categories for the student survey were: self-management (SM), self-efficacy in ELA (SE-ELA), self-efficacy in mathematics (SE-M), student perception of teacher (PT) and social-awareness (SA). The questions and their corresponding categories can be found in Appendix A. I also calculated a teacher perception of student (TPS) category using the teachers' responses from the "Teacher

report on students' MESH competencies" from the CORE District survey (2016) administered at the completion of this study in late February/ early March (see Appendix C).

The demographic variables of gender, looped status, and special education status presented statistically significant relationships with survey indicators. Gender significantly correlated with the SM indicator in December ($p = .009$), SA indicator in December ($p = .007$) and February/March ($p = .037$), and the TPS indicator ($p = .041$). Females reported higher levels of self-management ($M = 3.92$, $SD = .78$) and social-awareness (December, $M = 3.94$, $SD = .61$; February/March, $M = 3.97$, $SD = .68$) than males (December, $M = 3.58$, $SD = .79$; February/March, $M = 3.69$, $SD = .74$). Additionally, the teachers' perceptions of females ($M = 4.10$, $SD = .96$) were higher than for males ($M = 3.69$, $SD = 1.14$). The lone statistically significant relationship involving looped status was with the SM survey indicator ($p = .034$) in December. Students who looped scored themselves lower ($M = 3.55$, $SD = .97$) than those who did not loop ($M = 3.88$, $SD = .67$). Finally, Special Education enrollment presented a statistically significant relationship with the SM ($p = .010$), SE-ELA ($p = .026$), and PT ($p = .015$) indicators in December. Students in Special Education scored lower than their peers in self-management ($M = 3.24$, $SD = .93$) and self-efficacy in ELA ($M = 2.95$, $SD = 1.01$), and their perceptions of their teachers were lower ($M = 3.61$, $SD = .80$) than their peers (SM, $M = 3.83$, $SD = .77$; SE-ELA, $M = 3.53$, $SD = .90$; PT, $M = 4.07$, $SD = .64$) in December. There were no statistically significant differences in outcomes based on ELL status, FRL, or school enrollment. These results, including non-significant findings, can be found in Tables 21, 22 and 23 in Appendix H.

Phase I Quantitative Research Question 1: What are the relationships among students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness?

The student level data for this research question were analyzed and are presented in three stages. First, I present the means of the students' Acuity outcomes, their perceived levels of self-efficacy for each subject and their perceptions of their teachers' effectiveness, noting the changes in each. This is followed by an analysis of the relationships between students' achievement and the survey categories. Finally, I conclude this section with a report of the regression test results.

Students' Acuity outcomes increased from B to C in both ELA (December: $M = 460.5$, $SD = 31.29$; February: $M = 461.04$, $SD = 35.30$) and mathematics (December: $M = 446.28$, $SD = 29.71$; February: $M = 455.11$, $SD = 28.86$). This increase was only significant in mathematics ($p < .001$). Students' perceived self-efficacy as measured via survey showed a normal distribution for both ELA and mathematics for the December and February administrations. The mean self-efficacy scores during both administrations were higher for mathematics (December: $M = 3.66$, $SD = .94$; February: $M = 3.85$, $SD = .94$) than ELA (December: $M = 3.46$, $SD = .93$; February: $M = 3.60$, $SD = .89$) and the mean for both subjects increased from December to February. However, the increase was only statistically significant in mathematics ($p = .025$). The students' perception of their teachers' effectiveness also followed a normal distribution pattern for both December ($M = 4.02$, $SD = .67$) and February ($M = 4.13$, $SD = .67$), and presented an increase in the mean between survey administrations. This increase was not significant.

I examined the resulting continuous variables for relationships between each other and the student outcomes on Acuity B and C in each subject using correlational analysis. Student outcomes on Acuity B in ELA presented statistically significant correlations with only the SE-

ELA ($r = .259, n = 114, p = .005$) and TPS ($r = .211, n = 114, p = .024$) indicators. Meanwhile Acuity B mathematics outcomes correlated only to the TPS indicator ($r = .227, n = 114, p = < .015$). Student outcomes on Acuity C ELA and mathematics presented statistically significant relationships with SM, SE, and TPS (see Table 10). There was no significant relationship between students' achievement and their perception of their teacher in either administration window. However, all survey categories presented statistically significant relationships with each other in both windows. Tables 11 and 12 present these correlations.

Table 10

Results of a Correlation Analysis of the Students' Acuity C ELA and Mathematics Outcomes with the Significant Categorical Survey Indicators from the February/March Collection

Measure	1	2	3	4	5	6
1 – Acuity C ELA	-	.38***	.24**	.36***	.18*	.23*
2 – Acuity C Mathematics		-	.218*	-	.37***	.36***
3 – SM			-	.49***	.59***	.54***
4 – SE-ELA				-	.51***	.28**
5 – SE-M					-	.39***
6 – TPS						-

Note. SM = self-management. SE-ELA = self-efficacy in ELA. SE-M = self-efficacy in math. TPS = teacher perception of student. $N = 114$. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

Table 11

Results of a Correlation Analysis of the Categorical Survey Indicators from the December Collection

Measure	1	2	3	4	5	6
1 - SM	-	.58***	.59***	.34***	.55***	.44***
2 - SE-ELA		-	.61***	.22*	.51***	.24*
3 - SE-M			-	.28**	.59***	.34***
4 - PT				-	.40***	.34*
5 - SA					-	.38***
6 - TPS						-

Note. The TPS survey data was collected during the February/March window.

* $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

Table 12
Results of a Correlation Analysis of the Categorical Survey Indicators from the February/March Collection

Measure	1	2	3	4	5	6
1 - SM	-	.49***	.59***	.39***	.56***	.54***
2 - SE-ELA		-	.51***	.40***	.59***	.28**
3 - SE-M			-	.28**	.50***	.39***
4 - PT				-	.44***	.31**
5 - SA					-	.36***
6 - TPS						-

Note. ** $p < .01$, two-tailed. ***

Due to the significance of the SM, SE, and TPS indicators, I used hierarchical regression analyses to test if students' perceived self-efficacy and self-management levels and their teachers' perception of them predicted their Acuity C outcomes on ELA when controlling for looped status and special education enrollment variables. As presented in Table 13, results of the regression indicated SE-ELA explained 15% of the variance in ELA outcomes ($\beta = .31$, $p = .001$). SM ($p = .610$) and TPS ($p = .120$) variables were not significant predictors in the model.

Table 13
Predictors of Acuity C Outcomes in ELA

Variable	Model 1 B	Model 2	
		B	95% CI
Constant	-.25	-.09	[-.64, .45]
Not Looped	-.40*	-.35	[-.70, .01]
Non-SpEd	.57*	.36	[-.18, .89]
Self-Efficacy		.31**	[.13, .49]
R^2	.08	.17	
F	4.53***	7.37***	
ΔR^2		.09	
ΔF		12.14**	

Note. $N = 114$. CI = confidence interval. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

I ran a separate regression test to see if the SM, SE-M, and TPS variables predicted students' Acuity C outcomes in mathematics. As displayed in Table 14, results of the regression

indicated the SE-M ($\beta = .22, p = .019$) and TPS ($\beta = .27, p = .004$) predictors explained 23% of the variance in mathematics. SE-M alone predicted 17% of the variance in students' mathematics outcomes. SM was not a significant predictor ($p = .546$) in this model.

Table 14
Predictors of Acuity C Outcomes in Mathematics

Variable	Model 1 <i>B</i>	Model 2		Model 3	
		<i>B</i>	95% CI	<i>B</i>	95% CI
Constant	.17	.11	[-.64, .45]	.16	[-.35, .68]
Not Looped	-.60**	-.49**	[-.84, -.14]	-.52**	[-.86, -.18]
Non-SpEd	.24	.23	[-.28, .74]	.19	[-.30, .68]
Self-Efficacy		.32***	[.15, .49]	.22*	[.04, .40]
TPS				.26**	[.08, .44]
R^2	.09	.20		.23	
F	5.72**	8.88***		9.32***	
ΔR^2		.101		.06	
ΔF		12.86***		8.75**	

Note. $N = 114$. CI = confidence interval. TPS = Teacher Perception of Students

* $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

Phase II Research Question 2: What are the observable behaviors and norms of students in classrooms with cultures of varying levels of agency and self-efficacy (CASE)?

During phase II of my study I observed the behaviors of students and teachers in the 10 classrooms, utilizing grounded theory (Glaser & Strauss, 1967) to manually code and analyze my observation notes in a sequential exploratory process (Creswell & Plano, 2017; Creswell et al., 2003). After each round of observation and data collection, I utilized the resulting coding as a new version of the observation tool (see Appendix D and E for the first two drafts and Appendix F for the final tool). As discussed in Chapter 3, the iterative collection and analysis process allowed me to generate several versions prior to developing a final iteration of the observation tool used to collect quantitative data for triangulation. The quantitative results are discussed in subsequent sections. The remainder of this section describes the results of the inductive to

deductive process and contains descriptions and vignettes (Erickson, 1986) of the agentic and self-efficacious behaviors documented during classroom observations in Phase II that informed and help exemplify the indicators on the final observation tool.

Results of the research process. As explained in Chapter 3, the qualitative research process addressing this question was sequential cycle of collection, analysis, instrument-building, instrument-testing and instrument-revision. This took place over the course of four rounds. The purpose was to collect data on the agentic and efficacious student behaviors at the task-level, as research has demonstrated this has the strongest correlation with achievement (Bandura, 1997; Bong, 2001; Bong & Clark 1999; Bernacki, 2013; Foster, 2016; Lynch, 2013; Pajares, 1996, 2003). The following paragraphs explain the results from each round of observation.

Round one results. I initially organized my notes from my first round of 15-minute classroom observations according to three relational categories: teacher behaviors, student behaviors, and the learning environment. During the inductive analysis coding process, I did not use frequency as a significance indicator. Rather, I used a combination of Holistic Coding of lumped observable phenomena and Provisional Coding using my literature review for the language of the codes and guidance on appropriateness and relevancy of the data (Bloomberge & Volpe, 2016; Saldaña, 2013). Because my research question focused on student behaviors, the lens for the preliminary codes I derived was student rather than teacher-centric (Saldaña, 2013). These codes along with example student phenomena can be found in Table 15. I used these codes to develop the first iteration of the observation tool (see Appendix D for Tool 1).

Table 15
Open Codes with Example Data from One Round of 15-Minute Observations in the 10 Participating Classrooms

Code	Example Observation Data
Time management	Students took time to explore the details of the story to correct mistakes and misassumptions about the theme.
Transfer of knowledge	Student uses her prior knowledge of fraction equivalency to multiply $\frac{2}{3}$ with $\frac{3}{4}$.
Perseverance in problem solving	Students asked for help from peers and teachers rather than giving up when stuck on a problem.
Perseverance with solutions	Students engaged in a discussion justifying their answers to one another to the point of realizing their solutions were correct over others.
Set challenging goals	Students chose a task that was more challenging or chose to complete a task in a more challenging way than minimally required.
Celebrations	Students cheered and celebrated a student reaching his goal
Selection of efficient strategies	When given the option of creating a visual model to use to represent mathematics equations, students planned out their method prior to starting to avoid trial and error, and chose the most efficient model
Self-reflection	Students wrote reflections on their growth for the lesson stressing their role in their learning
Monitor understanding	Students rated their understanding on a 1-4 scale and had to give an example of the concept in a new context to demonstrate this level of understanding.
Monitor work habits	Students tracked their work effort and habits on charts displayed around the classroom.
Confidence in peer interaction	Students spoke with confidence during a discussion about reducing an answer to the lowest common denominator, using justifications to support their opinions (did not conform to peer pressure or dominate opinion)
Grade-level content	Students were given appropriate grade level content to work on or discuss as opposed to work that was too easy or below grade level.
Use tools of the content	Students incorporated academic vocabulary into their conversations and used the mathematical tools (arrays) to solve the multiplication problem.

Round two results. During my second round of observations using the first draft of the tool (see Appendix D), I initially attempted to document only the frequency of the coded student behaviors in the classroom. However, I struggled to do so in the context of the observation for a

several reasons, including not clearly defining what constituted an instance (a single student, a group of students, or the whole class) prior to observing. Therefore, I quickly switched back to exclusively inductive qualitative collection for this second round. During the analysis phase, I conducted Axial and Eclectic Coding (a combination of first cycle methods) on the new data refining my original codes, and then re-coded my original round of data according to new, final codes (Saldaña, 2013). These final codes were used in the final tool (see Appendix F for Tool 3) and are examined further in subsequent sections.

I also explored the relationships within and between the relational categories of student and teacher behaviors from both the first and second rounds of observation (Bloomberge & Volpe, 2016; Saldaña, 2013). I realized that my initial assumption of using a student-exclusive lens during the preliminary open coding process limited my data collection and also contributed to my struggle in this round of collection. Therefore, for this second round of analysis I also used Axial Coding (Saldaña, 2013) to identify the dimensional degree to which the student and/or the teacher owned the coded behavior. These were represented by a category of “ownership” codes: *Students Independently*, *Students with Support*, and *Teacher Independently*.

The ownership codes combined with the new observation codes were the basis for the indicators and collection methods on my second draft of the observation tool as presented in Appendix E. Tool 2 also attempted to address the limitations of the first draft by collecting data on the extent to which the ownership indicators were present in the classroom for each behavior indicator. I endeavored to address this through a Likert scale for each ownership indicator with the intent that it quantify and help triangulate my findings. However, the inclusion of three Likert scales for each indicator created another layer of complexity to the third round of observations that influenced the final iteration of the tool as explained in the next section.

Round three results. The last round of qualitative data collection and tool development was a combination of an inductive and deductive process complicated by the structure and measurement method of the second iteration of the tool. As explained in Chapter 3, I initially included a partner observer⁷ for this round to help establish inter-rater reliability on the tool, but quickly realized the Likert scales for each ownership indicator added unnecessary collection points, slowed the observation process, and made inter-rater reliability unlikely. Therefore, I quickly abandoned the idea that I would be able to collect valid quantifiable data during this round and continued to progress through the qualitative collection process used in round one without strict adherence to the tool.

During the third round of analysis I selectively recoded my categorical data into themes according to Bandura's (1997) sources of self-efficacy. This theming schematic became the organizing function for Tool 3 with the sources serving as the organizing source category subscales, the final behavior codes became the indicators, and the ownership codes became my measurement scale for each indicator. Specifically, the ownership codes became a Task Experience / Responsibility Scale consisting of a single four-point scale that allowed observers to document whether most students were exhibiting the behavior independent of the teacher (3) doing so with teacher support or prompting (2), whether the behavior was mostly the responsibility of the teacher (1) or if it was not observed (0). Figure 6 illustrates the product of this entire research process manifested in Tool 3.

⁷ My partner observer was an instructional coach in the secondary program at Garfield Academy who was also enrolled in a doctoral program and was in the dissertation proposal-drafting phase. She had no evaluative responsibilities for the teachers observed and had prior training in teacher observation and data collection both from the school, prior employers, and through her doctoral program.

Organizing Source Category Subscales

The organizing source category subscales for each indicator were derived from Bandura's (1997) sources of self-efficacy and identified during the round 3 analysis selective coding process.

Task Experience / Responsibility Scale:					
3 - Students independently of teacher					
2 - Students with teacher support or prompting					
1 - Mostly or all teacher					
NO – Not Observed					
Mastery Experiences	Experience Success	3	2	1	NO
	Think about / engage with rigorous, grade-appropriate content (do the cognitive work)	3	2	1	NO
	Utilize prior knowledge or strategies to persevere through work when challenges arise	3	2	1	NO
	Appraise and monitor progress	3	2	1	NO
Vicarious Experiences	Use peer models for guidance	3	2	1	NO
	Use peer models for self-evaluation or comparison	3	2	1	NO
	Respectfully share thinking and differing opinions	3	2	1	NO
	Interact with peers to support or complete work	3	2	1	NO
Verbal Persuasion	Critique work and give feedback toward the goal	3	2	1	NO
	Assess academic work	3	2	1	NO
	Assess work behaviors	3	2	1	NO
	Celebrate or offer encouragement for academic and/or CHOW accomplishments	3	2	1	NO
Physiological & Affective States	Regulate emotion when it comes to unsuccessful experiences	3	2	1	NO
	Regulate / manage behaviors	3	2	1	NO
	Manage time effectively / are productive	3	2	1	NO

Task Experience / Responsibility Scale

This scale was derived from the ownership codes originating in round two of the Axial Coding process.

Behavior Indicators

These indicators were derived from the behavior codes originating in round two of analysis via Eclectic Coding.

Figure 6. The result of the research process on the final iteration of the observation instrument, Tool 3, as presented in Appendix F.

Round four results. The final round of observations was purely deductive and was used to collect data using the final observation tool. In review of Chapter 3, prior to beginning this final round my observation partner and I convened to discuss, familiarize, and orient us to Tool 3 using the example data for the indicator codes from prior rounds of observations. Appendix G

contains the “look for” criteria document we used when training. My co-observer and I independently watched the same six video clips of classroom lessons to collect data using the tool and look-for criteria, and conducted four live classroom observations together to ascertain our degree of inter-rater reliability. I collected her independent responses and compared them to mine and found us to be in 80% agreement or better on each indicator (a 0.80 on Kappa’s Benchmark Scale). We were within one scale point of each other when not in agreement. I then conducted observations in all 10 classrooms in both ELA and mathematics using the final version of the instrument, Tool 3. My partner co-observed in all 10 classrooms with me for one subject that varied by classroom. The quantitative data collected during this process is presented under the results of Research Question 3 later in this chapter.

Observable behaviors and norms. The remainder of this section contains descriptions and vignettes of the observable behaviors and norms of students in classrooms with cultures of varying levels of agency and self-efficacy (CASE) as collected over the course of the qualitative Phase II process. According to Erickson (1986), the purpose of the inclusion of vignettes to support qualitative results is to ground the abstract in the concrete and to assist in validating the researcher’s interpretations. As explained previously, I used Holistic Coding of lumped observable phenomena in my field notes (Bloomberge & Volpe, 2016; Saldaña, 2013) and key portions of videos and informal student discussions were later manually transcribed and pictures cataloged for additional evidence and inclusion in these vignettes. Because of the sequential exploratory nature of the qualitative Phase II data collection, not all evidence or vignettes described in the following paragraphs were collected during the final, quantitative portion of the collection window. Therefore, any references to scoring are for the purpose of illustrating how the phenomena would have been scored had it been observed within the final round of the

February/March quantitative collection process. Means and standard deviations referenced for classrooms were included for contextually comparative purposes.

Mastery experiences. Using Bandura's (1997) description of enactive mastery experiences as subscale, I derived four indicators during the coding process that I used to collect data. Specifically, I attempted to capture the extent to which students: 1. *Experienced success*, 2. *Thought about / engaged with rigorous, grade-appropriate content (did the cognitive work)* 3. *Completed tasks that utilized prior knowledge or strategies to persevere when challenges arose*, and 4. *Appraised and monitored their progress*. The following paragraphs provide exemplars of each of the indicators organized under Enactive Mastery Experiences as scored on the Task Experience / Responsibility Scale.

Experienced success. Students who experienced success were those who effectively accomplished a task, completed steps in the process of working through a task, or successfully completed a component of the task after learning from prior errors. In some cases, evidence for the *experienced success* indicator was determined by the student and in others by the teacher's expectation for the class. For example, in many of the classrooms the teacher had posted learning targets (a standard written in student-friendly language) for the day's lesson either on the board or on work (see Appendix I for examples). Students in classrooms who reported high levels of student self-efficacy in informal discussions were observed getting correct answers on tasks aligned to the target.

In some cases, students experienced success working through the process of the task rather than achieving the stated outcome or learning target. Students in Classroom 2 with levels of academic achievement in mathematics in February above the mean ($M = .764$, $SD = .609$) and higher levels of student self-efficacy ($M = .697$, $SD = .735$) were asked to explain whether or not

a stated quotient to a solved mathematics problem was reasonable or not without actually going through the process of computing or solving it. The stated target was “I can demonstrate my understanding of division by solving multi-step real world problems” but no students actually solved a division problem during the lesson. Instead, students participated in small group discussions at their tables about fractions as division problems demonstrating problem solving skills and conceptual understanding about the relationship between division and fractions. This was the exchange between students:

Student A: “So... is that why it is called the denominator? ‘Cause when we divide [with] fractions you are, like, just multiplying [by] the denominators...”

Student B: “Yea, but that’s not it. [The denominator] tells you how many pieces the whole is broken into... so there are more of them.”

Student C: “Uhhh... So I don’t think she solved it right because her... [quotient] is a smaller number, not a bigger one.”

In this case, students demonstrated a conceptual understanding of what division was and its relationship to fractions, and concluded successfully that the quotient of the given problem was not reasonably correct. These students were experiencing success while not necessarily successfully solving division problems as the target intended. Additionally, several students in other groups who initially concluded that the problem was correct demonstrated success with the content as they corrected their misconceptions with other problems over the course of the observation. This lesson is an example of a 3 due to the lack of teacher guidance or intervention in the success of the students’ achievements toward the intended learning target.

Conversely, lessons were observed in classrooms with lower levels of perceived student self-efficacy where students were generally unsuccessful or the teacher was the only one who

experienced success. In Classroom 8 where the participating students' mathematics achievement on Acuity C ($M = -.131$, $SD = .917$) and perceived self-efficacy in mathematics ($M = -.207$, $SD = .982$) were lower than the mean, the mathematics target on the board was "I can sketch polygons by reading the descriptors". At the start of the observation students were working independently on a worksheet. The teacher circulated about the room prompting the whole class to work a bit faster and addressing mistakes she was seeing. These comments ranged from "If you get stuck, work on" and "Hurry up; don't forget about the back" to "Be careful on number six. Is it part-part-whole? I see a lot of people doing 6×24 , that's not right." Upon observation only seven students completed the problems with some level of success (either by process or by computation to a right answer).

During a whole group review of the answers the teacher walked the class through her process of drawing the polygons based on the descriptions. Her comments included, "First, I drew a rectangle... Then we labeled the side lengths... Next, we wrote a multiplication sentence... Last, we multiplied to find area." A few students corrected their work but most sat and watched her work, including most of those who had it incorrect on their papers. Only one of the four students with whom I conversed indicated they felt successful on the task and had met the learning target. This lesson exemplified a score of 1 for this indicator.

Rigorous, cognitive work. Bandura (1997) noted that being assigned unchallenging tasks leads to lower self-efficacy levels. Therefore, for the *thought about / engaged with rigorous, grade-appropriate content (did the cognitive work)* indicator, I looked for evidence that the work the students were asked to complete was at or above grade level. In mathematics indicators included the aspect of rigor of assigned work, and in ELA whether the text used was appropriately complex for the grade level (at or above grade level as measured by Lexile or

qualitative level). In both subjects the extent to which all students were engaged in work that was an appropriate Depth of Knowledge (DOK), typically Level 2 or above, was also used to determine whether the work was sufficiently rigorous. DOK was a construct developed by Webb (1997) frequently utilized in education to classify the depth of cognitive complexity and thinking required to perform a task. Level 1 is defined as the recall or reproduction of skills or knowledge (Webb, 2002). This level is characterized by questions where the student is either right or wrong and does not have to engage in solving or analyzing the concept. Conversely, Levels 2 – 4 require students transform known information and skills before answering questions, and are differentiated from each other by the number of pieces of information and thinking processes utilized in the solution (Webb, 2002).

As discussed for the *experience success* indicator, most teachers posted learning targets in their rooms or on student work. Observers compared the written targets to the grade level standards, the appropriate DOK for the standard, and the work students were asked to complete. If targets were not present, observers used the student tasks as evidence and examined them for alignment to the corresponding standard(s) and required DOK levels. In cases where there was alignment with the grade level standard and DOK, the lesson could score a 1, 2, or 3 based on who had ownership over the cognitive work. If the target or task was below grade level grade level or DOK was inappropriately a Level 1, the indicator was automatically scored as Not Observed (NO).

While the division questions in the prior example for Classroom 2 described under the *experienced success* indicator required students to have right or wrong answers, the emphasis of the lesson was not on students working towards correct computations. Rather, students spent the majority of the observed portion of the lesson synthesizing knowledge about division and

fractions to apply them in new situations both at and above grade level. The target was “I can demonstrate my understanding of division by solving multi-step real world problems” and, as demonstrated in the prior vignette, students were engaged in work and discussions of mathematical problems aligned to Indiana standards 5.AT.4, as well as 6.C.4⁸ (a year above grade level). Therefore, this lesson exemplified a score of 3 for this indicator.

In Classroom 7 where participating students statistically scored at the mean on their perceived self-efficacy in ELA in February ($M = .001$, $SD = 1.107$) but below the average on the Acuity C ELA assessment ($M = -.139$, $SD = 1.032$), I observed the teacher asking low level (DOK 1) recall questions to the whole group of students. Here is an excerpt from the ELA lesson:

Teacher: There was one other thing they learned how to do. What was it Alex?... Fish, cook, clean...?

Alex: (unintelligible)

Teacher: Not quite, you are skipping ahead a little bit. What else did they learn to do at winter camp Mark? Take care of themselves and their little....

Student (not Mark): [shouting out] Babies!

Teacher: Children! They were learning to take care of their younger siblings. Right? So they were learning about more of those life skills that we talked about. Were they learning about arithmetic or math?

Students: [shouting out] No!

Teacher: Were they learning how to read?

⁸ The language of 5.AT.4 is “Solve real-world problems involving division of unit fractions by non-zero whole numbers, and division of whole numbers by unit fractions (e.g., by using visual fraction models and equations to represent the problem)” (IDOE, 2017). The language of 6.C.4 is “Compute quotients of positive fractions and solve real-world problems involving division of fractions by fractions. Use a visual fraction model and/or equation to represent these calculations” (IDOE, 2017).

Students: [shouting out] No!

Teacher: How to write?

Students: [shouting out] No!

Teacher: No, they were learning more of those life skills. Alex skipped ahead to chapter 10. Go to page 65. We spent a lot of time talking about a word on there. Concertina.

We started talking about concertina and we started talking about context clues... we look at other words before and after, we look at the chapter title or heading, we always look back to those things to help us find the meaning of unknown words. So what was a concertina? Lets see if you remember what that is.

Student: an accordion

Teacher: an accordion!... What do accordions do?

Student: They play music.

Teacher: They play music. They used that to celebrate at the end of winter camp which is what Alex was talking about.

The learning target for this lesson was “I can describe characters and explain how their actions contribute to the sequence of events in a story” which was aligned to Indiana standard 3.RL.2.3. However, as demonstrated in the prior excerpt, questions observed during the lesson were not related to the target and asked students to recall events in the story and the meaning of words (the child who answered “accordion” did not reference the text); skills not aligned to grade level standards. Furthermore, the teacher frequently prompted students for right answers rather than for analysis or synthesis of the information or concepts, exemplifying a score of Not Observed (NO) for this indicator.

Prior knowledge or strategies to persevere. For the *completed tasks that utilized prior knowledge or strategies to persevere when challenges arose* indicator, I looked for students utilizing strategies such as sounding out unknown words, utilizing computers, dictionaries, manipulatives, or other tools to problem solve, relying on their conceptual understanding to figure out a mathematical procedure, etc. I also looked for students using other soft skills such as conflict resolution, organization, etc. to problem solve obstacles to complete their tasks.

An observation in Classroom 3 exemplified a score of 3 for this indicator. Participating students in this classroom performed slightly above average on Acuity C mathematics ($M = .025$, $SD = .817$) as well as reported above average mathematics self-efficacy levels in February ($M = .197$, $SD = 1.401$). The target for this lesson was “I can create symmetrical patterns”. Students were observed working in groups creating a pattern, image, etc. using tiles of various shapes and lines drawn on their desks (see Figure 7 for examples). Students utilized prior knowledge of lines of symmetry for simple shapes to create more complex ones with varying degrees of success. The teacher’s role during this lesson was observer and questioner, as she did not prompt students to use any specific strategies or give any specific guidance on how to make corrections.

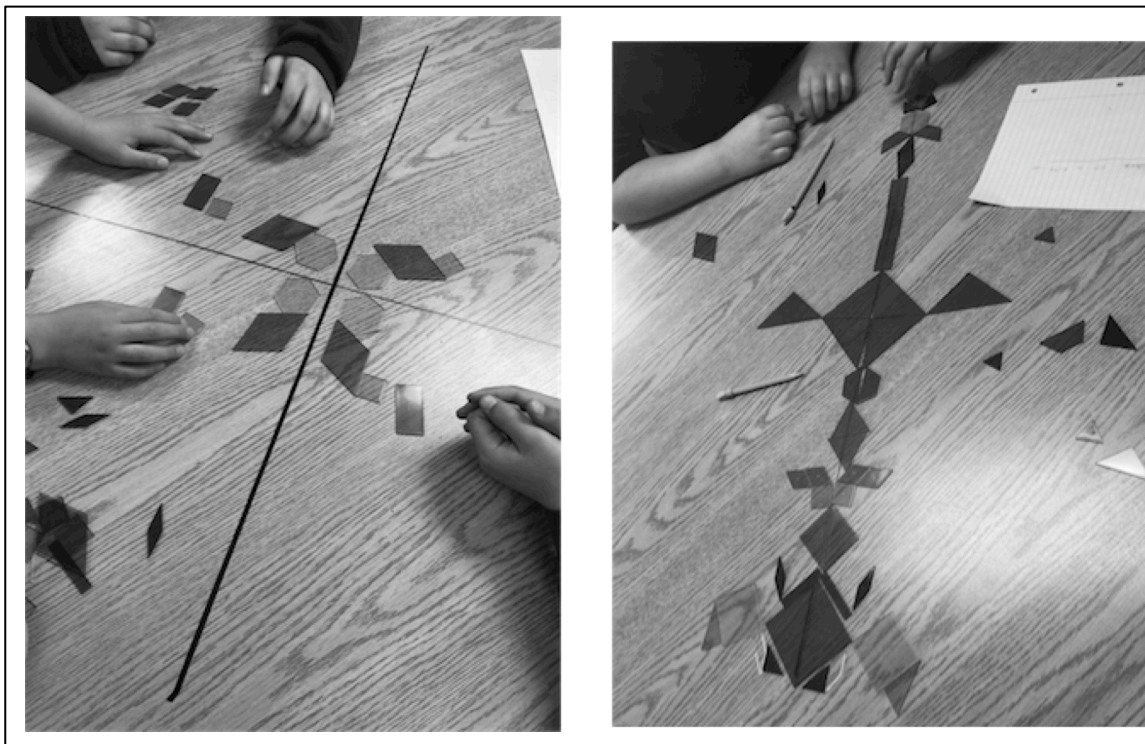


Figure 7. Examples of symmetrical patterns created by groups of students in Classroom 3. The image on the left is of a group who progressed through a single line of symmetry (such as the image on the right) and began working on creating patterns utilizing multiple lines of symmetry.

Many groups were observed utilizing innovative tools and strategies to help them determine the correct orientation of some of the tiles to re-establish the symmetry of their figures. I observed one group initially unsuccessful at creating symmetrical patterns using a mirror from a students' makeup compact to create a mirror image of the pattern along the line of symmetry. After some conversation about what the mirror was showing them, the students in the group were able to shift the tiles appropriately to make the figure symmetrical. Another group utilized paper to make impressions of the raised tile shapes on one side of their line, flipping it over the line of symmetry as a mirror image to compare it to the other side of their figure to make adjustments to the tiles.

All of the students' behaviors described in this observation of Classroom 3 were unprompted and unregulated by the teacher. Conversely, participating students in Classroom 9 scored below average on mathematics Acuity C ($M = -.527$, $SD = 1.155$) but above the mean for perceived self-efficacy in mathematics in February ($M = .110$, $SD = .608$). During an observation of Classroom 9, students were asked to create symmetrical snowflakes from pipe cleaners then trace them onto paper, labeling the angle measurements after measuring them with a protractor. While the teacher frequently verbally prompted students to use the protractor and stopped to model it for two individual students, the majority of students were utilizing it as a ruler rather than a tool to measure angles to ensure they were identical. Furthermore, no students utilized the protractor to create their initial pipe cleaner snowflakes, making their tracings asymmetrical from the start. As a result, most students were unable to produce a symmetrical shape and did not receive the feedback that they were utilizing the tool incorrectly. Students were given general feedback that their figures or angle measurements were wrong, and grew increasingly frustrated resulting in all but one participating student opting out of completing the task. Contrary to the example from Classroom 3 where students utilized tools and their knowledge of symmetry to help them complete the task, students in Classroom 9 were not able to utilize prior knowledge of a protractor and did not utilize information about symmetry to persevere. Therefore, this lesson typified a score of NO.

Appraise and monitor progress. The last indicator for the Enactive Mastery Experiences subscale measured students' ability and opportunity to *appraise and monitor their progress*. For this indicator, I looked for the extent to which students actively self-assessed their progress towards their success on the task, targets, or behavior objectives, and to what extent they did it independent of the teacher, with teacher modeling or support for understanding the expectations,

if it was done to the student without their active participation, or if it was absent from the classroom during the observation period.

This indicator was well exemplified in several classrooms at a score of 3. Participating students in Classroom 7 scored above the mean in mathematics on Acuity C ($M = .464$, $SD = 1.107$) and mathematics self-efficacy ($M = .217$, $SD = .783$). They were observed turning in their mathematics exit ticket into one of three folders on the wall, “Got It!”, “Almost!” and “Nope!” without prompting from the teacher (see Figure 8). Students told me they regularly self-select in which leveled folder to place their exit tickets based on their perceived progress toward the task’s learning target. The lack of teacher prompting beyond providing the leveled folders also indicated this was a regular, normed practice for the students.

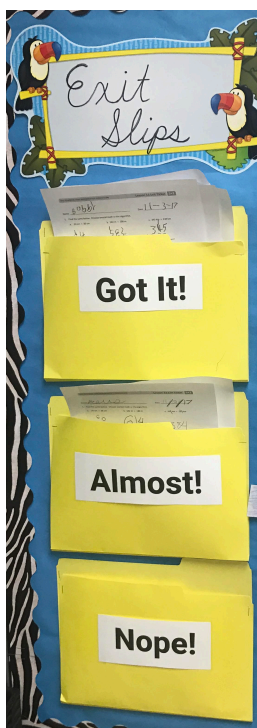


Figure 8. Students in Classroom 7 turned their mathematics work into leveled folders posted on a side bulletin board based on their perceived level of progress towards their success on the task.

Similarly, students in other classrooms turned their completed work into leveled bins without teacher prompting (see Figure 9). In Classroom 3, Level 1 was labeled “Beginning - I am starting to get it, but I am still confused. I am just starting to learn this but don’t understand it completely.” Level 2 was labeled “Developing – I can do this with help or an example in front of me. I kind of get it but may make a mistake”. Level 3 was labeled “Meeting – I can do this on my own without help. I can show I understand.” Finally, Level 4 was labeled “Exceeding – I can do this on my own. I can explain how to do it. I can teach someone else how to do it.”



Figure 9. Students in several classrooms turned their work into leveled bins students, such as this one found in Classroom 2.

Students in Classroom 2 not only turned their work into these bins without teacher prompting, they frequently documented their perceived level of achievement on their work and provided a written reflection on why they scored themselves that way. Figure 10 is an image of a single piece of student work from a larger collection hung on a bulletin board. At the bottom of the task, it shows the student’s self-assessment circled at a level 3 and contains an explanation for that ranking. The student wrote, “I feel confident in every way. I check it and work it out step by step.” I used this as evidence that students actively self-assessed their progress towards their success on the task, targets, or behavior objectives.

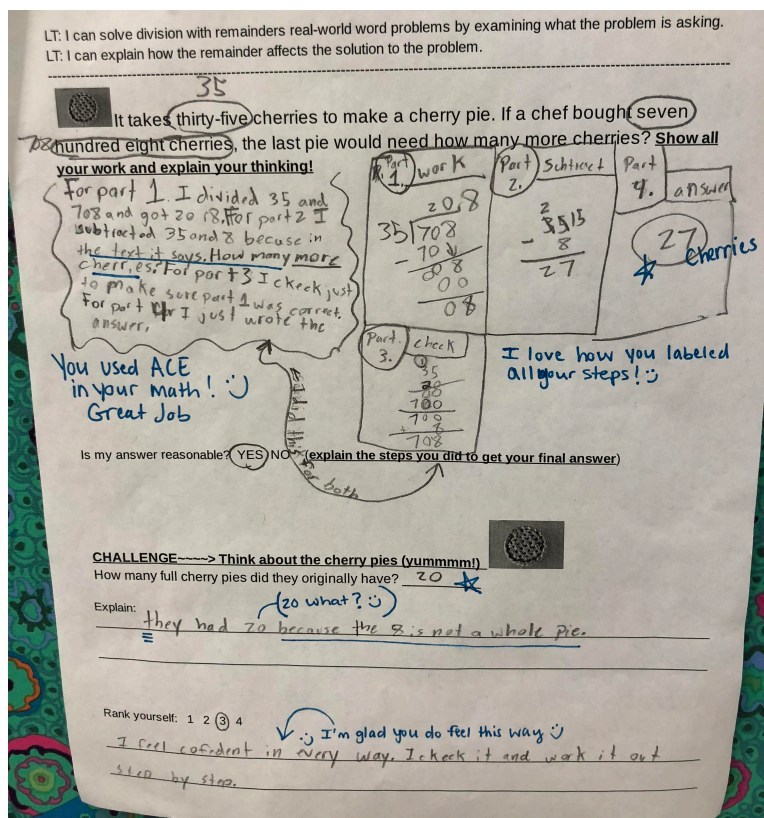


Figure 10. Student work hung on the bulletin board showing students' self assessment and justification for the score.

Students were not provided with the same opportunities to self-assess their progress in all classrooms. Participating students in Classroom 1 performed below average on both the Acuity C mathematics assessment ($M = -.655$, $SD = .778$) and perceived self-efficacy in mathematics ($M = -.865$, $SD = 1.293$) and were rarely asked to self-monitor or appraise their progress towards success during the observed mathematics lessons. During all of the mathematics observations conducted in this room, the teacher attempted to ask students to self-appraise only once. During this attempt, the teacher asked students to put their thumb up if they were ready to move on to independent practice. However, the teacher disregarded that almost all students signaled they were ready to move on and continued to model two more problems for the class. Additionally,

the teacher routinely made the appraisal and monitoring observations herself such as “John is correct” and “That is too loud. One more warning, sir.” Because the teacher determined the students’ progress towards successful task completion, learning, and the behavior goals, this example was representative of a score of 1 for this indicator.

Vicarious experiences. In review, Bandura (1997) described vicarious experiences as the opportunities for students to appraise their proficiencies and capabilities in relation to peer models. Therefore, I found four aligned indicators through the Phase II sequential collection and coding processes that I then used to collect data during the quantitative collection window. I used the Task Experience / Responsibility Scale again to differentiate the degree to which students were exhibiting these behaviors with teacher intervention or not. Specifically, I attempted to capture the extent to which students: 1. *Used peer models for guidance*, 2. *Used peer models for self-evaluation or comparison*, 3. *Respectfully shared their thinking and differing opinions*, and 4. *Interacted with peers to support or complete work*. The following paragraphs provide examples of each of the indicators organized under the Vicarious Experiences source subscale as scored on the Task Experience / Responsibility Scale.

Use peer models for guidance. For the *use peer models for guidance* indicator, I collected evidence on the extent to which students looked to other students or their work for guidance on what to do, how to solve a problem, how to revise their thinking, how to behave, etc., or what Bandura (1997) documents as “strategy knowledge” (p. 223). In the prior example of Classroom 3 described under the *prior knowledge or self-regulation strategies to persevere* indicator where students created patterns along the lines of symmetry, students were frequently observed visiting other groups to look at their work. Students made comments to one another

about complexity and uniqueness of the symmetrical tile patterns. One group created an owl, which prompted other groups to attempt to create their own unique animals.

Students were also observed using their peers as models for problem solving during this lesson. The student's compact mirror mentioned previously began being passed around the room when groups realized how it could help them determine the correct orientation for misplaced tiles. Furthermore, as groups finished and saw other groups making more complex patterns with additional lines of symmetry, they began to draw their own additional lines without being prompted to do so by the teacher. In each case, groups were observed creating unique patterns, indicating they were using their peers' models to develop strategy knowledge rather than simply copying others' work.

In contrast, students in Classroom 5 scored slightly below the mean on the Acuity C mathematics assessment ($M = -.071$, $SD = .864$) and perceived self-efficacy in mathematics ($M = -.316$, $SD = .764$). During one observation the students worked at tables on comparisons and conversions of numbers to the hundredths and their equivalent fractions. While a few students wrapped up a worksheet, the majority of students used decimal cards to write equivalent fractions on their tables with dry erase markers (see Figure 11). I did not observe students use each other as resources or peer models. Several students wrote incorrect fractions for their decimal numbers. Others incorrectly aligned the whole number with the numerator (see Figure 11). These students were also not observed using their peers to guide corrections or revisions to their work. Instead, the teacher moved from table to table, providing students the model for how to interpret their decimals and produce equivalent fractions. One exchange between the teacher and a student went like this:

Teacher: This is the tenths place and this is the hundredths [pointing to each digit on the card]. So that is twelve hundredths [writing the fraction $\frac{12}{100}$ on the table]. What about something with a whole number in front of it? [Hands student a new fraction card.]

Student: Mixed fraction.

Teacher: Yea, a mixed fraction.

Student: [Writes $1\frac{27}{100}$ on the table.]

Teacher: Yea... So I would suggest making your one a bit bigger just so it's clear that it's the whole number [draws over the one in the whole number place to make it extend to the top and bottom of the adjoining fraction rather than aligning with just the numerator]. So go ahead and grab your marker and get started.

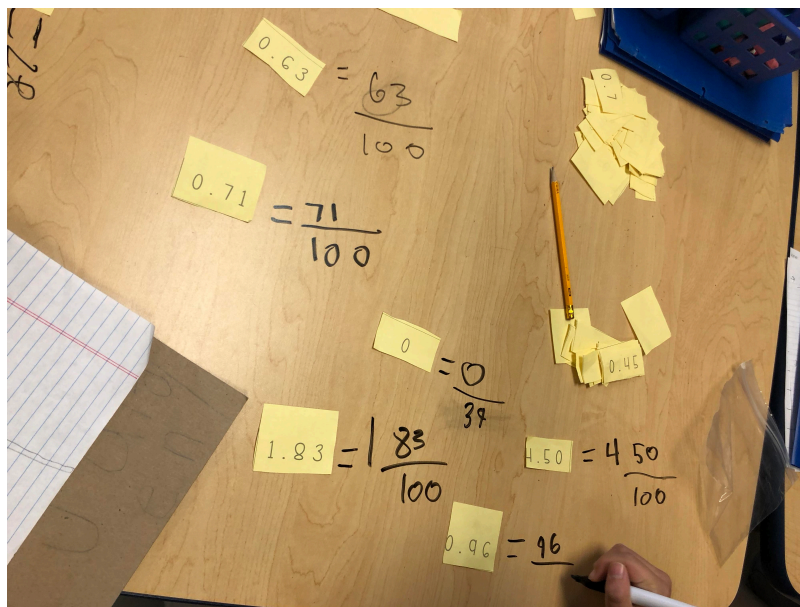


Figure 11. Students in Classroom 5 write equivalent fractions for decimal numbers on their desks using dry erase markers. This student struggled with writing whole numbers next to the fraction to clearly indicate its status as a whole number, independent from the numerator.

The students at the table began working independently while the teacher moved to a new table to model equivalent fractions again. However, the student the teacher worked with from the prior example began to make mistakes when she encountered numbers with just tenths. She wrote .6 as $\frac{6}{100}$ using the pattern she had discussed with the teacher. Rather than use her peers for guidance when I asked her about it, she struggled for a bit to explain why she wrote it that way and raised her hand to ask the teacher for help. The student continued to use the teacher as a model for copying the pattern rather than developing her strategy knowledge. In this case the teacher acted as the sole model for students on the task, exemplifying a score of 1 for this indicator.

Use peer models for self-evaluation or comparison. This indicator required evidence that students not just use each other for guidance on their work, but that they actively compared their work to determine their level of mastery or success in comparison to their peer. This indicator scored a 3 when students not only used each other's work as a model but also engaged in making comparisons. An observation of ELA in Classroom 1 best exemplified this indicator at this level.

Students in Classroom 1 scored slightly above the mean on the Acuity C ELA assessment ($M = .050$, $SD = 1.176$) and perceived self-efficacy in ELA in February ($M = .370$, $SD = 1.172$). During the observation students were engaged in a discussion about their peer's writing. The learning target for the lesson was "I can use quotes and citations to support the claims in my writing." The student, Jan, was at the front of the room with her written work under the document camera. She read the entire piece aloud and her peers engaged in conversations about its qualities relative to the learning target (the details are discussed later in the *Critique work and give feedback toward the goal* section). Eventually, the teacher asked the class to determine the

level of mastery Jan demonstrated over the learning target based on this evidence. The students and teacher agreed that the piece demonstrated Jan had mastered the use of using textual quotes and citations to support the claims in her work. Then students were asked to rate their level of mastery over their own use of quotes and citations in pencil at the top of their drafts using Jan's as a model for mastery, and were instructed to begin the revision process on their own pieces. Because students actively engaged in comparing and assessing what was quality about the peer exemplar and the degree to which it compared to their own work, this lesson exemplified a score of 3 for this indicator.

Several classrooms used generic public data charts and graphs to compare students' level of development towards various classroom objectives. These documented everything from success on Acuity, to their implementation of the RACE (Restate, Answer, Cite, Explain) structure for responding to a question or writing prompt, to tracking their behavior / CHOW (Character and Habits Of Work) in class. CHOW is a metric used by both schools for students and teachers to discuss and track seven specific student behaviors, centered on respect, responsibility, integrity, independence, collaboration, goal-setting, and academic grit (see Appendix J for the specific targets and rubric). Examples of these public data charts and graphs can be found in Appendix K. The following example shows how using these as simple peer comparison points better describes a score of 1 for this indicator.

Students were observed engaging with the teacher tracking their Acuity data in Classroom 6. Participating students in this classroom performed above average on both of the Acuity C assessments (ELA: $M = .787$, $SD = .955$; Mathematics: $M = .299$, $SD = .813$) and perceived self-efficacy (ELA: $M = .344$, $SD = .692$; Mathematics: $.214$, $SD = .899$). As students prepared for their student-led conferences at the end of the quarter in March, they conferenced

with the teacher one-on-one regarding their Acuity C benchmark achievement percentages. The teacher shared the students' overall percent correct on the assessment and had them graph it, marking their quarterly goals and class average on the assessments as well (see Figure 12). While this gave students some insight into how well they were performing relative to their peers and their goals, the conversation was teacher-dominated and not specific to "what" they were succeeding or struggling with. In the two conversations I observed, only one student talked and it was to ask about the placement of the class average marking. Furthermore, data on this tracking sheet largely served as a generic data point for students rather than, as Bandura (1997) notes is effective practice for building self-efficacy, including a specific and qualitative point of reference for what they were doing well on or needed to improve upon in each subject. For these reasons, this portion of the observation exemplified a 1 for this indicator.

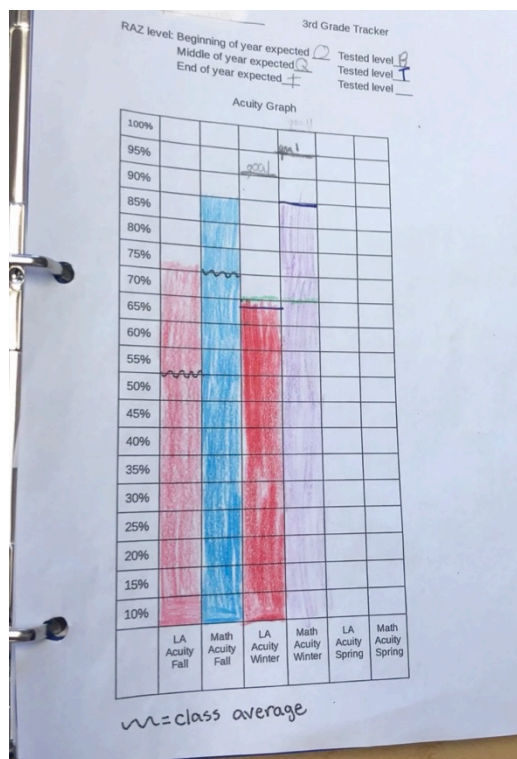


Figure 12. Students tracked their achievement on the Acuity assessments relative to their goals and the class average in a bar graph.

Respectfully share thinking and differing opinions. Bandura (1997) noted, “students must learn how to select and structure environmental settings in ways that are conducive to learning” (p. 228) and added that students are more likely to experience a positive learning environment if they are considerate of their peers. Therefore, I included this indicator to capture the extent to which students collaborated and shared their thinking and feedback with each other respectfully, disagreed without anger or conflict, and the tone of their general interactions with each other to structure a positive environment conducive to learning. Classrooms where there was an established culture where students respectfully shared and engaged in conversations about their thinking with each other on academic subjects without teacher moderation or control scored a 3 for this indicator. Classrooms where the teacher had a heavy hand in moderating or controlling

the conversation and/or the way students heard or interpreted each other's thoughts were indicative of a score of 1.

During an observation of mathematics in Classroom 2, students were sitting at tables in groups of three and four engaged in conversations with each other about a series of problems on a mathematics worksheet. While the teacher conversed with individual groups asking questions such as "Why did you choose that strategy?", the students not working with the teacher were observed engaging in respectful conversations with each other. In general, students were overheard saying things like, "I would like to add on to what ____ and ____ was [sic] saying..." and "I disagreed with my group but they convinced me that we should divide". The tone of all students observed was respectful even during disagreements on mathematical processes and answers. For example, I overheard the following conversation during a debate about whether the whole number in a fractional answer was two or three:

Student 1: "Now, what were you saying about reducing?"

Student 2: (unintelligible)

Student 3: "... it would be two and 19 over 15 but that is an improper fraction.

Student 1: "But how did you get three?"

Student 3: Because this is an improper [fraction] and you have to reduce that. And so you have three and you are left over with $\frac{4}{15}$.

Student 1: Yea! Because 15 from 19... (unintelligible)... and its $\frac{4}{15}$. Thank you! So now we have to write sentences explaining how and why...

While the previous vignette was an example of an observation that would score a 3 for this indicator, not all classrooms had consistent evidence that students regularly interacted respectfully with each other and developed environments that were as positive and conducive to

learning. For example, when I walked into Classroom 9 for an observation during their scheduled reading block, students were sitting in their chairs in a circle. Rather than discussing the text students were holding, they were conversing about a behavior incident that had just occurred. A student I talked with revealed the discussion began after several students accused five boys of throwing objects during the read-aloud of the text and the boys got upset that the other students were “tattling”. During the observed discussion, the teacher called on several students to speak up about how it felt to get hit with the objects. However, students began calling each other names and accusing each other of being bullies. The teacher tried summarizing and reframing the students’ comments to include only the parts about their feelings, but participating students continued to struggle to have a respectful conversation without interference and moderation from the teacher. Eventually, the teacher shut down the conversation and directed students to read the story quietly at their tables. This observation, while not entirely academic in nature, was representative of a score of 1.

Interact with peers to support or complete work. Bandura (1997) noted that self-efficacious learners use the resources around them to learn and succeed at tasks. This includes peer resources. The intent of the *interact with peers to support or complete work* indicator was to capture the extent to which students are free to use peers when and how they want to complete the task or learn content, and take the opportunity to do so. Evidence for this indicator overlapped with several others previously explored (see Classroom 1 under the *respectfully share thinking and differing opinions* indicator, and Classroom 3 under the *use peer models for guidance* indicator) as observed classrooms were outfitted with tables rather than individual desks and students frequently had the freedom to rely on their peers for support. However, some

classrooms had more subtle evidence of students' use of peers to learn and complete tasks. One such instance of this, illustrating a 3 for this indicator, is explained in the following paragraphs.

Although I did not observe direct implementation in Classroom 3, there were student exemplars on the board during one observation (see Figure 13). At the top of each paper were the following learning targets with a score⁹ next to each: "I can multiply a 4 digit number by a 1 digit number using partial products" and "I can justify my answer as well as if someone else's answer is correct or incorrect [sic]". Not all answers from the papers displayed on the board were perfect; some had common mistakes highlighted yellow.

⁹ The grading scales at the two schools represented a continuum of development, rather than the tradition A-F system, with 1 being a Beginning level, 2 was Developing, 3 was Accomplished or Meets, and 4 was Exceeds or Exemplary/Leader.

Partial Products Quiz:

LT: I can multiply a 4 digit number by a 1 digit number using partial products. 3
 PS.3: I can justify my answer as well as if someone else's answer is correct or incorrect. 3

3,906 x 3

Partial Products:

3,906
 x 3

 11,718

Area Model:

3,000	900	6
x 3		
9,000	2,700	18

2,450 kids participate in the spring soccer program and 4 times that many kids participate in the fall soccer program. How many kids participate in the fall?

Partial Products:

2,450
 x 4

 9,800

Area Model:

2,000	400	50
x 4		
8,000	1,600	200

Every batch of cookies Jasmine made needed 3 cups of flour. She wanted to make 396 batches of cookies for her school bake sale. How many cups of flour will she need?

Partial Products:

396
 x 3

 1,188

Area Model:

300	90	6
x 3		
900	270	18

Figure 13. Evidence from Classroom 3 of an exemplar quiz displayed on the front board identifying the computational error in yellow, and checking the correct use of methods and computation for the remaining problems.

In the work presented in Figure 13, the student earned a 3 for each target. The student made a computation error in the first question and the mistakes were highlighted in yellow, circled and commented upon by the teacher. However, the student answered the remaining questions correctly. When I asked a student why the quizzes were posted and what they were used for, she mentioned that the teacher regularly posted exemplars of work from students in the class. Students were allowed to redo work or retake an assessment on their learning target only after they looked at the exemplars and submitted a written response to the teacher about the

mistakes they made and their new learning, and felt ready to take another assessment. This opportunity to redo work was student-initiated and not teacher-mandated.

When I asked the student how this impacted her learning and her belief about her ability in mathematics she responded:

It helps me know what to do... This one time I didn't realize I was multiplying decimals wrong, like putting the decimal in [the wrong] place. I was so mad when I got a one. But I looked at [another student's] paper - he got a four on that target - and I still didn't understand [what I was doing wrong] but I went to him and he showed me how to do it... I had to write [to the teacher] to explain what I had been doing wrong and [show] how I knew to fix it... As Mrs. ____ always say [sic], "Its not that we can't do it, we just can't do it *yet*".

The student continued to explain that she had her work from another unit posted on the wall previously as well. She said she was eventually able to earn a 4 on her learning target as a result because she helped teach her friend during their work time, and her friend got a 3 on her second attempt at the assessment. This was strong evidence that students in Classroom 3 have the opportunity to work with their peers when and how they want to complete the task or learn content, and take the opportunity to do so.

There were several examples of students working in isolation with a lack a peer support during classroom observations. As mentioned previously in the example of Classroom 8 where students were working on sketching polygons by reading the descriptions of various shapes, the students worked completely alone and did not collaborate with peers on their construction. Rather, the teacher was the only point of interaction for students and even then, not all students interacted with her to learn the content. There also was no other evidence, such as student work

on the walls as in the prior example, that students regularly had the freedom to work with their peers and use them as resources during this observation. Therefore, this was an example of a NO for this indicator.

Verbal Persuasion. Bandura (1997) described verbal persuasion as affirmation of one's abilities to achieve a goal. It is derived through persuasive and affirmative feedback, suggestions and guides for improvement when goals are not met, support of self-management strategies such as self-appraisal and self-improvement, and celebratory encouragement for acquiring new knowledge or skills. Therefore, I found four aligned indicators through the Phase II sequential collection and coding processes that I then used to collect data during the quantitative collection window. I used the Task Experience / Responsibility Scale again to differentiate the degree to which students were exhibiting these behaviors with teacher intervention or not. Specifically, I attempted to capture the extent to which students: 1. *Critiqued work and gave feedback toward the goal*, 2. *Assessed academic work*, 3. *Assessed work behaviors*, and 4. *Celebrated or offered encouragement for academic accomplishments and/or hard work*. The following paragraphs provide examples of each of the indicators organized under the Verbal Persuasion source subscale as scored on the Task Experience / Responsibility Scale.

Critique work and give feedback toward the goal. Bandura (1997) frequently stressed the importance for students to receive feedback, and noted the impact of student self-appraisal on building and sustaining self-efficacy. It was this self-appraisal of competency and capabilities towards goals, short or long term, that defined this indicator. Specifically, I looked for evidence that students reflected on the quality of their work as it related to the learning target, and/or gave feedback or used steps in feedback protocols to reflect on their progress towards meeting their learning targets, goals, CHOW targets, etc.

Some classrooms used rubrics and conversation starters and released students to provide feedback to one another on concepts and learning targets (see Figure 14). Other teachers planned for whole-class critiques of a student's work and achievement toward a learning target. For example, returning to the observation in ELA in Classroom 1 where students were using their peer's (I will call her Jan) work to assess their own, the teacher asked the class to give Jan feedback on how well she used quotations and citations from the text to support her claims in each paragraph. The students engaged in a conversation about the quality of work, noting where she had two strong citations in one paragraph but had none in another. The class eventually determined this was acceptable because it was a transition rather than an explanatory paragraph. While initially the teacher prompted the conversation to focus on Jan's use of citations, the students almost entirely ran the discussion about the quality of her work from this lens. Furthermore, when a student made a comment about her nice handwriting adding clarity to the piece, another student reminded the class to focus their feedback, and proceeded to comment on Jan's helpful use of paragraph numbers in her citations to further specify evidence referenced from the text. The students then engaged in the revision process and self-assessed their own progress towards the learning target. This observation was an example of a 3 for this indicator.

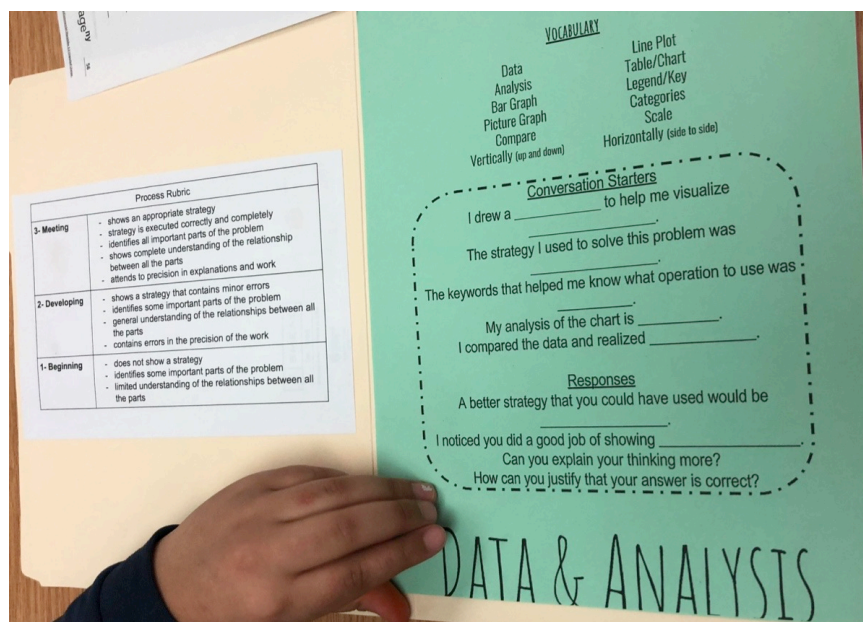


Figure 14. Students in some classrooms were provided with rubrics and conversation starters to help provided targeted feedback and critique to their peers. In this image the mathematics process rubric is on the left and the conversation starters along with specific mathematics vocabulary is on the right. Students used these to give one-on one feedback to one another.

Occasionally, the teacher was the only person to give students feedback toward their learning targets during the course of the observation. In Classroom 4, the students' achievements on Acuity C in both subjects were slightly above the mean (ELA: $M = .015$, $SD = .879$; Mathematics: $M = .164$, $SD = 1.093$) while their perceived self-efficacy levels were below the mean in both subjects (ELA: $M = -.719$, $DS = 1.212$; Mathematics: $M = -.150$, $DS = 1.198$). During an observation in Classroom 4, students were given standards-based progress reports and completed growth goals sheets (this process is examined more fully in the *Assess academic work* indicator). The progress reports listed the significant learning targets for the quarter and noted the expected level of development next to the individual student's grade (see Figure 15). Areas of strength were highlighted in yellow in ELA, Mathematics, and CHOW while areas for

improvement were highlighted in orange. This is the only feedback students engaged with during this observation and therefore represents a score of 1 for this indicator.

English Language Arts				
My Level	Expected Level			
57	50%			
Reading Acuity Score				
Reading Literature				
I can identify a story's point of view.				
3	3			
I can read and comprehend 4th grade appropriate fiction text.				
3	3			
I can make inferences about characters' traits based on the character's thoughts in the text.				
3	3			
I can conduct research on a specific topic, using reliable sources.				
3	3			
Writing				
I can use RACE to answer constructed response question with evidence from the text.				
4	3			
I can write a hook for a narrative by establishing the situation and introducing the characters.				
3	3			
I can organize my narrative using HCAST.				
3	3			
I can develop the characters by using dialogue and describing how they react to situations.				
3	3			
I can use transitional words and phrases to help my narrative flow.				
3	3			
I can incorporate concrete details in my narrative to make my writing more descriptive.				
3	3			
I can use proper grammar, spelling, capitalization, and punctuation in my narrative.				
3	3			
Social Studies				
Indiana History				
My Level	Expected Level			
I can identify where I live including my continent, country, region, state, and city.				
3	3			
I can describe native tribes of Indiana (Native Americans).				
3	3			
Mathematics				
Math Acuity Score	My Level	Expected Level		
48	50%			
Division Unit				
I can represent and solve division problems with up to a 4-digit dividend.				
3	3			
I can use models to solve real world division problems with a remainder.				
3	3			
I can demonstrate my understanding of division by solving multi-step real world problems.				
2	3			
Review				
I can round to a given place value.				
3	3			
I can compare and order multi-digit numbers.				
3	3			
I can identify factor pairs for a given number.				
3	3			
I can add and subtract multi-digit numbers.				
3	3			
I can multiply two 2-digit numbers and a 4-digit by a 1-digit.				
3	3			
I can solve long division problems using a standard algorithm and/or a model.				
3	3			
I can find the area and perimeter of rectangles and irregular shapes.				
2	3			
Fractions				
I can write, identify, and model equivalent fractions.				
3	3			
I can compare fractions using common denominators, common numerators, or benchmark fractions.				
3	3			
I can add and subtract fractions.				
2	3			
I can add and subtract mixed numbers.				
3	3			
Attendance				
Q1	Q2	Q3	Q4	
Days Absent	0	0	0	
Days Tardy	0	0	0	
CHOW				
Q1	Q2	Q3	Q4	
I take responsibility for my work, learning, and actions. This means I also own what I do and say, when I both do well and make mistakes.				
3	3	3		
I am respectful to adults and my peers. This means my words, actions, and tone of voice are polite, appropriate for school, and represent CHA well.				
3	3	2		
I am independent. This means I can think, act, and work in my own when appropriate.				
3	3	3		
I have integrity. This means I am honest and do what is right even when it is hard.				
3	3	3		
I am collaborative. This means I actively listen to others, contribute to ideas and conversations, and work with others to solve problems and create good work.				
3	3	3		
I have grit. This means I think through problems and don't give up when things are hard.				
3	3	3		
I set goals, take action to meet them, and reflect on my progress.				
3	3	2		

Figure 15. A sample student progress report from Classroom 4. The learning targets were listed under each subject with the student's current scores for each and the expected level of development, along with the most recent Acuity score. Areas of strength were highlighted in yellow in ELA, Mathematics, and CHOW while areas for improvement were highlighted in orange.

Assess academic work. In his work, Bandura (1997) frequently noted the impact of self-monitoring and goal setting on cognitive development. Setting and meeting goals helps fulfill a student's sense of self-accomplishment thereby building and sustaining their sense of self-efficacy. Therefore, under the *assess academic work* indicator, I looked for the extent to which students looked for right and wrong "answers" or their successful application of skills within their task or work, the extent to which they made value or quality judgments about the outcomes of their own work, and whether they set goals for their academic achievement.

One such class, Classroom 6, exemplified a 3 for this indicator in several different observations. During an observation of ELA in December the teacher engaged the students in a reflection of their academic behaviors (described in more detail later in the *assess work behaviors* section) and then transitioned students to reading individually with their own text. She reminded them to reorient themselves to their self-selected reading goals posted on the wall (see Figure 16) and sent them to find a reading space in the room. Students immediately began reading, choosing to read silently or quietly out loud to themselves. Many were doing a combination of reading and jotting down notes in their reading journals. Looking through students' journals, it was apparent the self-monitoring process was a consistent practice in the classroom and that there did not appear to be a specific format or process for it. One student read silently but would pause to write words with page and paragraph numbers next to them in her notebook. I noted that some words had stars next to them. When I asked about her annotation strategy, she remarked that she was documenting all of the words she did not know but was using context clues to figure out. The stars meant she could not use the context clues to determine the meaning of the words and needed to look them up. Another participant told me she did not want to engage in discussion at that point in the class because her reading behavior goal was to read the whole time.

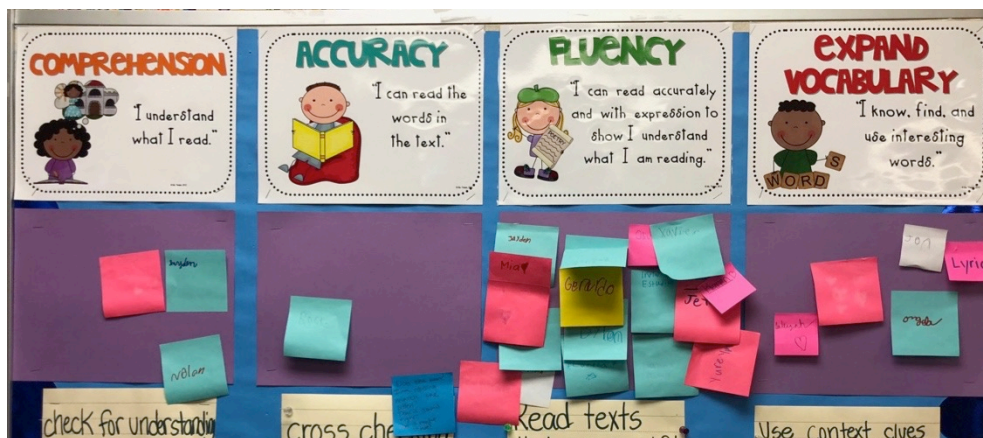


Figure 16. The learning goals for comprehension, accuracy, fluency, and expand vocabulary were listed on posters on the wall in the front of Classroom 6. Students had written their names on the front of a post-it and their specific strategy on the back for how they were going to meet their self-selected goal.

I conversed with a student in Classroom 6 who chose to sit on the rug in the front of the room and who was engaged in reading out loud quietly without taking notes. He explained that he set a goal for improving comprehension through fluency after reading one-on-one with the teacher the week prior. He said,

I really want to work on saying words with expression, like, reading exclamation marks... [Last week,] if it had a question mark, I wasn't reading them like that - reading them as questions. [Because I do not always notice punctuation] I get confused about what [characters] say and stuff as I'm reading out loud sometimes.

He then showed me his post-it with his name on the front and strategy on the back located under the fluency target on the wall, and explained that he would do a self-reflection about his progress towards his goal in his journal after reading.

Furthermore, during an observation in Classroom 6 in March, the students were observed assembling work in their portfolios and writing goals during their literacy block. For the assembly process, they were directed to choose pieces of completed work that showed their growth and/or accomplishments for their subjects and to choose other, less accomplished evidence that justified the improvement goals they set for themselves for the final quarter of the year. The goals ranged in subjects from writing to music, but the students were fairly consistent in their use of evidence to support their targeted subject. While it was clear that the process was teacher-directed and students were prompted to reflect, they already had the skills required to do so and were doing it at a fairly consistent level with one another. Multiple students were heard asking each other what evidence they were using or what subject they were setting goals in, and a group of girls engaged in conversations about the merits of using multiple drafts of a writing piece to show their accomplishments in ELA. A boy at a neighboring table inserted himself into the conversation and convinced them to include the drafts. He remarked that he was including his drafts as well as the final copy as justification for an improvement goal in writing because he noticed his comma omissions while the final draft hung in the hallway. Verbally he revealed he wanted to try to use more dialogue in his writing to make it more interesting, although he left it at “using punctuation to make my writing more interesting” (see Figure 17). Because students were largely engaged with self-appraisal, assessment, and goal-setting without major support from the teacher Classroom 6 consistently represented a score of 3 for this indicator.

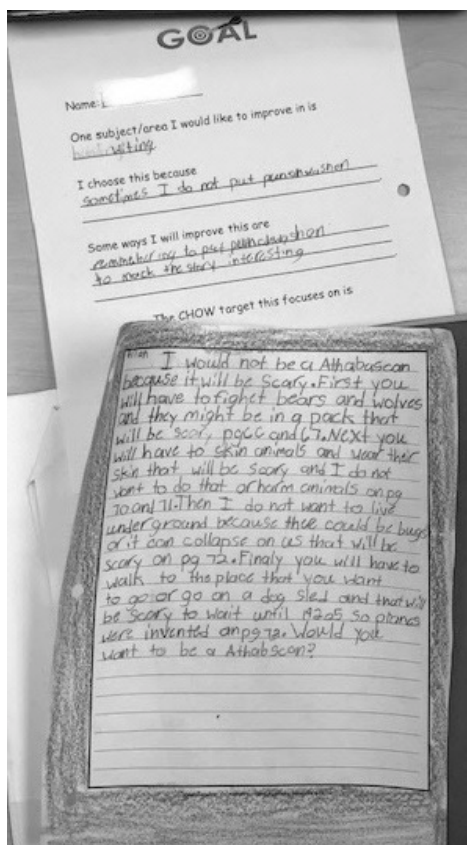


Figure 17. A student's draft of his goal sheet for his portfolio along with his drafts and final copy of his corresponding writing piece containing errors with commas.

Students in Classroom 4 also developed goals but they were much more teacher-directed. During the previously mentioned observation where students were given their highlighted, standards-based progress reports (see Figure 15), students were directed to complete growth plans based on the information on the reports. For ELA, mathematics, and CHOW, the teacher had highlighted areas of strength in yellow while areas for growth were highlighted in orange. To complete their growth plans, students were assigned to write goals for the upcoming quarter using the targets highlighted in orange. Students wrote their targets on goal sheets along with their current score and their target score (see Figure 18). Whereas the students in Classroom 3 were asked to reflect on their development in a standard before being allowed to reassess (see the

section on the *interact with peers to support or complete work* indicator), students in Classroom 4 were asked to reflect after re-taking an assessment and used the growth plan sheets to help track which targets they were working towards continuing to develop. However, because students did not make value or quality judgments about their own work and the teacher essentially gave students their goals for academic achievement, this type of assessment of academic work exemplified a 1 for this indicator.

Growth Plan:

Learning Target I would like to improve: I can represent and solve division problems with to a 4-digit dividend.

Current Score: 1 Goal Score: 3 Score I received: _____

What didn't you understand that you understand now about this learning target? _____

Figure 18. Students in Classroom 4 used these growth plan sheets using the highlight goals from their progress reports to set goals for the next quarter.

Assess work behaviors. Bandura (1997) cautioned against using effort as a reinforcement lever for promoting self-efficacy, as it can actually have the reverse effect if not utilized effectively in conjunction with teaching explicit strategies helpful for converting efforts into successes. The schools use CHOW (Character and Habits of Work) targets to help define those specific work behaviors for students (see Appendix J for the targets and rubric) without a specific “effort” category. Therefore, this indicator sought to capture the extent to which students rated, scored, or otherwise assessed themselves on those specific work behaviors/CHOW targets, and/or established behavior/CHOW goals.

Eight of the 10 participating classrooms had sticker charts posted for each student to publically track their character and habits of work accomplishments (see Figure 19). In general, students in these classrooms stated that they received stickers when they exemplified specific behaviors for the CHOW target in class. Students in four of these classrooms consistently indicated they self-tracked and put stickers on their own charts, students in the other four responded that they mostly received stickers from the teacher for demonstrating their accomplishment of the specific target. Only one of the eight classrooms, Classroom 2, had students stated that they also gave stickers to each other if they recognized a peer exemplifying the behaviors listed on the chart. Students in the two classrooms who did not have a public tracking system were observed to have an informal, task-based appraisal system.

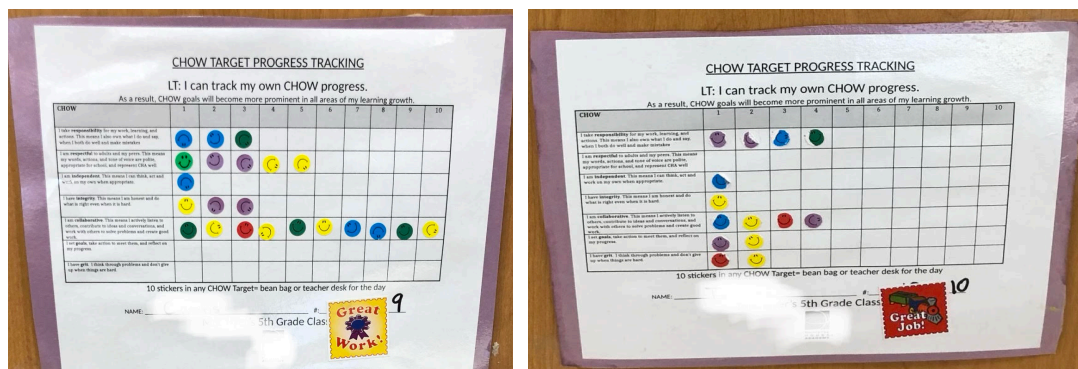


Figure 19. CHOW tracking sticker charts from Classroom 2. Students in this classroom indicated they were able to give each other stickers in recognition of the CHOW behaviors that contributed to their academic success.

In one such classroom with a more informal system, Classroom 6, students were observed checking for and monitoring their behavior objectives over the course of several observations. During the previously mentioned observation in December (see the *academic work* indicator), students were asked to do a quick personal check of their behaviors after working with a partner. The teacher asked students to sit facing her on the rug with their hand in

a fist on their chests. Students were instructed to put a thumb up if they felt like they met the behavior expectation, a thumb to the side if they had some work to do, and a thumbs down if they really needed to work on it. Here was the exchange between the teacher and her students:

Teacher: We need to make sure we are being responsible and having integrity with our reading behaviors with our partners so we are going to do a quick check.... Get ready for your personal check with your fist at your heart. [Teacher reads from a hand written poster of the behaviors on an easel.] Everyone was reading with someone...

Students: [Silently and with eyes closed, all put a thumbs up on their chest.]

Teacher: If you were sitting elbow, elbow or knee to knee... This is a personal reflection.

Students: [Most students put a thumbs up, two with thumbs to the side, and four with thumbs down].

Teacher: If you were using a level one voice...

Students: [Most put thumbs up, two put thumbs down, three put thumbs to the side.]

Teacher: Were you and your partner reading the whole time or is using your time wisely something you need to work on?

Students: [About half put thumbs up and half put thumbs to the side.]

Teacher: Remember, this is a personal check. Were you and your partner sitting in your spot the whole time?

Students: [All students put a thumbs up.]

Teacher: Did you get started right away, you and your partner, or is that something you need to work on?

Students: [Most had thumbs up, four had thumbs to the side. One student whispers something intelligible to the teacher.]

Teacher: That's okay if it's to the side guys. That gives us something to work on.

[Teacher points to and references the poster.] Were you looking at the words as your partner was reading or is that something you need to work on?

Students: [Most had thumbs up.]

Teacher: Check for understanding... Did you make a good deal about what you were reading and how you were dividing the work?

Students: [Mix of thumbs up, to the side, and down.]

In classroom 6, there was evidence that students rated, scored, or otherwise assessed themselves on work behaviors/CHOW and set behavior/CHOW goals for themselves and therefore exemplified a 3 on this indicator. However, those classrooms where the teacher rewarded decontextualized effort, simply told students that they were presenting work behaviors/CHOW targets, and/or set goals for the class exemplified a 1 on this indicator. I observed this during an ELA lesson in Classroom 5 as described in the proceeding vignette.

Participating students in Classroom 5 scored just above the mean on Acuity C in ELA ($M = .125$, $SD = .896$) and right at the mean for ELA self-efficacy ($M = .001$, $SD = .806$). During an observation of a small group literacy activity, the teacher walked around and issued students stickers for demonstrating positive behaviors primarily in independence and collaboration. The following lists the teacher's comments to the students during the course of the observation related to this indicator:

- I see (student) managing her time well. She's already on [question] four. Nice work staying on task. Good integrity and independence.
- (Student) is also doing a great job managing his time... (Unintelligible) Good job!

- Thank you (student) for using eye contact to show (student) you are listening. [Hands student a sticker.] Go put this on collaboration.
- Table three [is] demonstrating great collaboration. They are working together, everyone is on the same question... they are cooperating with each other.
- (Student), you are distracted. Put that away. What [question] is your group on?... I expect better.
- Stay focused! Don't give up! You are almost there. Rely on your group if you're stuck.
- Good job (student). Much better than yesterday... You good?... Yeah, keep going.
- I like how (student) is using all of the resources to help the group figure out [question] five. He is referencing [the text]. (Unintelligible)... Good problem solving, group.

Upon stating the above comments and compliments, the teacher issued stickers to the appropriate students and directed them to place them on their CHOW charts displayed on the wall at the side of the room. Because only the teacher assessed student work behaviors, this observation was representative of a score of 1 for this indicator.

Celebrate or offer encouragement for academic and/or CHOW accomplishments.

Whereas I looked for formal assessment processes for setting goals and assessing academic and behavioral accomplishments for the *assess academic work* and *assessing work behaviors* indicators, for the *celebrate or offer encouragement for academic and/or CHOW accomplishments* indicator I looked for less formal celebratory activities. These included giving compliments, cheering, saying “good job”, etc. when students made progress or had instances of doing good work.

Classroom 2 had strong evidence of a score of 3 for this indicator. Not only did I hear students say “good job” and offer compliments to one another, particularly in mathematics observations where there tended to be more collaborative activities and peer-to-peer interactions, but students also had a distinct ritual of offering complements to one another. During several observations students gave “shout outs” to other students in front of the whole class. Typically, students started with “I want to give a shout out to...” and issued a compliment to their peer. I noted students gave other students these “shout outs” for not giving up when solving a hard problem (four instances), helping another student without being asked (two instances), having a good idea during group work, and continuing to work hard to get work completed even though the student didn’t feel well. In all of these cases, the students initiated the “shout out” and gave the compliment in front of the class without prompting by the teacher.

Several other classrooms also had clear rituals the students engaged in when they noticed their peers excelling or progressing. Students in Classroom 1 regularly gave a “roller coaster cheer” where all students motioned their hands up then swopped down and finished with a single clap. Students in Classrooms 8 and 10 “kissed their brains” when they had a correct answer or were on the right track towards a right answer. In this case, individual students were called out to kiss their hands and touched their foreheads (or, kiss their brains) for getting correct answers. Unlike in Classroom 2, the teachers initiated all of the observed celebratory activities in these three classrooms. Students did not initiate or prompt a single observed instance. Therefore, these activities exemplified a score of 1 for this indicator.

Furthermore, the vignette described from Classroom 5 under the *assess work behaviors* indicator was also indicative of a score of 1 for this indicator. The teacher was the only person observed explicitly celebrating or encouraging either academic or behavioral accomplishments.

No students engaged in giving compliments or encouraging words to one another. Even when students were given stickers, and removed themselves from the group to place them on the wall, the other students in the group continued to work and none were observed congratulating or encouraging their peers for demonstrating the effective work behaviors. This also was indicative of a score of 1 for this indicator.

Physiological and Affective States. Bandura (1997) described the physiological and affective states as the impact the body and emotions play in our functions and reactions to situations. Particularly powerful was how we react to and manage stress and failure, both physically and emotionally. While Bandura recognized the relationship mood (on a continuum of happy to depressive) has with self-efficacy beliefs, I found this too hard to observe objectively without conversing with each individual student. Therefore, I aligned three indicators through the Phase II sequential collection and coding processes that I then used to collect data during the quantitative collection window. I used the Task Experience / Responsibility Scale again to differentiate the degree to which students were exhibiting these behaviors with teacher intervention or not. Specifically, I attempted to capture the extent to which students: 1. *Regulated emotion when it came to unsuccessful experiences*, 2. *Regulated/managed their behaviors*, and 3. *Managed their time effectively / were productive*. The following paragraphs provide examples of each of the indicators organized under the Physiological and Affective States source subscale as scored on the Task Experience / Responsibility Scale.

Regulate emotion when it comes to unsuccessful experiences. Bandura (1997) described the importance of students becoming adept at self-regulating their emotions and thoughts, particularly when it came to unsuccessful experiences. Therefore, I attempted to capture the extent of these abilities under the *regulate emotion when it comes to unsuccessful experiences*

indicator. Fully expecting students to encounter challenges in their academic work, I specifically looked for students' and teachers' reactions to these barriers to success in classrooms. To score a 3 for this indicator, I looked for evidence that when students were unsuccessful the majority either did not get frustrated or became frustrated but then self-regulated to continue persevering through the task. A score of 2 was issued if a teacher interceded to help the student and co-regulated, whereas a score of 1 was reserved for instances where teachers provided major support, such as redirection, intervention to remove the barrier / frustration, or they took the responsibility for regulation away from the student (such as a timeout from class or modifying the task). A score of NO was reserved for cases where there was no evidence that the task created frustration, confusion, etc. or no evidence that the students were able to handle it if present (such as students refusing to work, etc.) even with teacher intervention. The following examples help illustrate these distinctions.

The observation of the symmetry lesson in mathematics in Classroom 3 described under the *using prior knowledge or strategies to persevere* indicator also provided several examples of students utilizing self-regulation strategies to persevere when the work became too challenging. For example, I observed one student working on an additionally complex pattern using double lines of symmetry becoming seemingly frustrated, walking away from his table for a moment. He took a noticeable deep breath, observed another group for a moment and returned to his group, successfully re-integrating himself back into the work. Many students in several groups became frustrated over the course of the observation and in every case but one, the teacher was not required to intervene to help students self-regulate.

Conversely, I observed an instance of teacher-regulation and task modification in mathematics class (indicative of a score of 1) in Classroom 10. Students in Classroom 10 scored

below the mean on Acuity C in mathematics ($M = -.531$, $SD = .922$) and in self-efficacy in mathematics ($M = -.207$, $SD = .982$). Students worked on a multi-step mathematics problem over the course of the observation. The problem read, “James is saving money to buy a \$425 snowboard. He has \$245 and gets a \$9 weekly allowance. How many weeks will it take James to save enough money to buy the snowboard?” After asking a student to read the problem aloud, the teacher defined the words “allowance” and “snowboard” for the class. She then asked students to identify what they would do first and called on several students to respond. Each responded with a different answer that essentially involved variations of dividing \$425 or \$245 by \$9. One student thought they should add \$425 and \$245. During this period of interaction the teacher responded with the following comments:

- “No”
- “Be confident in your answer. Do you want a friend to confirm whether or not you’re right?”
- “Be confident, I’m asking you.”
- “If you think that’s what we’re doing, tell me... We’re not dividing it and splitting it into equal groups.
- “If anyone has anything different I want to know so they can figure it out.”
- “You think we are adding it?”

Interestingly, in this observation, not only did the students begin to get frustrated, the teacher appeared to also get frustrated by their responses. Her words became shorter, she spoke louder, and she began to pace in front of the board. She eventually stopped asking the students questions and began solving the problem for them. She set up the subtraction problem of “\$425 - \$245” on the board and asked a student to come to the board to solve it. Once the student had

modeled the difference on the board, she asked the class what operation they would use to find out how long it would take James to raise the remaining \$180. Students again began to answer incorrectly and many students began to have side conversations about subjects other than mathematics. The teacher complimented those students who were continuing to focus and not give up, and then set the division problem up for the students on the board. The students quieted down and copied " $180 \div 9$ " onto their papers. The teacher solved " $18 \div 9 = 2$ " then identified that she "dropped down" the 0 to get the dividend of 20 as the answer to the problem and asked a student to add the label of "weeks" to complete the response. She then asked the class to copy the work and put the response into their own words on their paper. The teacher eventually wrote the sentence, "It will take James 20 weeks to save enough money for the snowboard" onto the board, effectively removing the requirement that the students put the answer into their own words. In this observation, the teacher had modified the task to the point of removing all barriers and sources of frustration for students, and is, therefore, representative of a score of 1 for this indicator.

Regulate / manage behaviors. The *regulate and manage behaviors* indicator measured the extent to which students regulated their academic behaviors toward tasks (such as working at an appropriate pace, staying on task, appropriately working alone or with peers, etc.) without needing reminders from the teacher. Instances indicative of a 3 permitted students to receive reminders from their peers, whereas classrooms that required intensive intervention from a teacher for students to regulate and manage their behaviors (such as changing group work to individual work to get students to stay on task, isolating students from group work, removing a student from the classroom, etc.) were scored as a 1. The following paragraphs provide examples for this indicator.

There were many cases illustrative of a score of 3 for *regulate and manage behaviors*. During the lines of symmetry lesson observed in Classroom 3 as described in the *prior knowledge or strategies to persevere* indicator, students worked and managed their behaviors independent of the teacher for almost the entirety of the observation. In this case, students were provided a large amount of freedom in their academic behaviors. They were free to interact with peers within and between groups, access and utilize tools needed for success, move on or slow down depending on their success with the task. In these cases, the academic behaviors of students were not regulated by the teacher, but were influenced by peers.

Students observed during an ELA lesson in Classroom 4 had much less freedom by comparison, but still were representative of a 3 for this indicator. While students sat quietly listening to an audio recording of their text, *Lunch Money* (Clements & Selznick, 2007), they were observed watching the words being highlighted on the screen as they were read aloud by the narrator, following along in their books, and/or writing notes about the text in their notebooks. All students were engaged in academic behaviors that would help them comprehend the text they were reading. No students needed reminders from the teacher regarding the academic behaviors they should have been engaged in during the observation.

Conversely, during one observation in mathematics in Classroom 1, several students were off task and did not work on the mathematics problems unless the teacher stood next to them. After two boys engaged in an argument over a coin she later confiscated, the teacher responded to one of the boys with “Find your spot. You need to show some grit, integrity, and responsibility.” However, the students’ behavior changed only temporarily and they continued to struggle to self-regulate and manage their behavior. None of the students struggling with their academic behaviors received any further or more instructive feedback on corrective actions or

how their new behavior should look, but they were eventually separated from the group and scattered about the room to work individually on their assignment. Most of the separated students proceeded to work on the task. This lesson was representative of a 1 for this indicator.

Manage time effectively / are productive. The *manage time effectively / are productive* indicator intended to measure the extent to which students were observed working at an appropriate pace for the task, where time was not wasted, nor where students worked so fast that they made silly errors or took shortcuts in their work. There were several observations where the students were required to manage their time and demonstrated effective productivity, such as in the symmetry observation in Classroom 3, the division lesson in Classroom 2, and in the ELA lesson in Classroom 6. Students in these observations made adequate progress towards learning the content and completing the tasks without wasting time or taking short cuts to complete the work in the given time frame. These instances were indicative of a score of 3 for this indicator.

Several teachers were observed using timers and other time-tracking tools to help keep the class on pace. However, teachers and their students' responses to those varied widely from class to class. For example, during the observation of Classroom 2 as described under the *respectfully share thinking and differing opinions* indicator, students were provided with a two minute warning from the teacher before they began a whole-group debrief of the problem set. However, the teacher modified her timeline based on the conversations she overheard and work she saw from students. She announced, "I'm hearing great conversation... [and] see that most of you need more time to finish. I think three more minutes ought to do it. I'll check back [in] then." From the initial two minute warning to when the students were finally pulled together for the debrief activity, students remained working on the problem set. Those groups that finished early were observed reviewing their work and checking their answers with other groups. While

the teacher managed the students' time from a macro level, there was evidence that she was responsive to the needs of students, letting their management on the task dictate to some extent her lesson pacing.

Teachers in other classrooms were observed using timers and other time-tracking tools to more forcibly manage students' time. Students in Classroom 8 scored below the mean in ELA on Acuity C ($M = -.335$, $SD = .916$) and above the mean in ELA self-efficacy ($M = .120$, $SD = .852$). During an observation in ELA, the students were engaged in an activity requiring them to answer questions about a text they just read. The teacher utilized a countdown timer to help students track how much time was left until they transitioned to the next activity. During the observation the teacher routinely announced how many minutes were left and which question the students should be working on. For example, she announced, "There are four minutes left. You should be on question five. If you aren't, you need to work faster to [get to] question five." She made comments similar to those in her mathematics lesson described under the *experience success* indicator, such as "hurry up", "you're behind", and "keep going".

As a result of the teacher's actions and comments, several students were observed skipping ahead to question five to stay on pace. Other students quickly wrote down answers to questions three and/or four and hurried on to start on question five. Several of these students also appeared to rush through question five to get to the final question within the allotted time. Other students just gave up on the task. When asked why, one student responded, "I'm not [going to] finish it anyway." Another commented, "I hate the timer. I never finish [on time] and I just [feel] stupid."

When the timer went off the teacher collected the papers. Upon quick observation, it was clear that most students did not finish the task and some had only partial answers for some

questions. Because the students were largely not responsible for managing their behavior, and the teacher micromanaged the student behaviors to the point where students were largely inefficient and unproductive, this is an example of a lesson that would score a NO for this indicator.

Phase II Research Question 3: What relationship do students' agentic and efficacious classroom behaviors have with their perceived self-efficacy levels and academic achievement?

As described in more detail in Chapter 3 and previously in this chapter, during the inductive to deductive sequential explanatory collection process of Phase II (Bloomberge & Volpe, 2016; Creswell, 1999; Creswell & Plano Clark, 2017), I used rounds of observations and informal discussions with students to create and refine an observation instrument. In review, I organized the final version by source of self-efficacy (mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states), supported by specific, observable behavior indicators, and data were collected during the final observation round using a scale for "ownership" conditions I titled Task Experience / Responsibility Scale (see Appendix F for the final iteration of the tool). Rather than simply collect the number of instances of observed phenomena during an observation, I created a single four-point scale that allowed observers to document whether most students were exhibiting the behavior independent of the teacher (3) doing so with teacher support or prompting (2), whether the behavior was mostly the responsibility of the teacher (1) or if it was not observed (0).

I observed 10 classrooms for these 15 indicators in both ELA and Mathematics. In order to conduct the following analyses I coded classrooms as singular cases and attached the students' mean Acuity outcomes and survey data by classroom to the appropriate case. I conducted a

correlational analysis of the category subscales with each other. All subscales in ELA presented a statistically significant relationship ($p < .05$) with each other, as presented in Table 16. In mathematics, only the VE and PAS correlated ($r = .733$, $n = 10$, $p = .016$).

Table 16
Results of a Correlation Analysis of the Observation Category Subscales in ELA

Measure	1	2	3	4
1 - ME	-	.89**	.86**	.83**
2 - VE		-	.95***	.76*
3 - VP			-	.71*
4 - PAS				-

Note. ME = mastery experiences. VE = vicarious experiences. VP = verbal persuasion. PAS = physiological and affective states. $N = 10$. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

I did not have a large enough sample size ($n = 10$) from which to conduct a factor analysis on the observation indicators within each subscale. Therefore, I used Cronbach's alpha to calculate the internal consistency of the indicators in my observation tool, and to determine reliability within the organizing source category subscales. The specific indicators for each category subscale were previously identified in Figure 6 and are identified in the final version of the observation tool located in Appendix F. I conducted a subscale analysis for ELA and mathematics separately as well as combined, and analyzed the reliability of the collective tool within each subject. With the exception of verbal persuasion in mathematics ($\alpha = -.14$), all items indicated acceptable categorical subscale reliability within and between subject areas ($\alpha > .6$). The results are displayed in Table 17.

Table 17

Reliability of CASE Observation Indicators Overall and as Subscale Categories in ELA, Mathematics and Combined as Measured by Cronbach's Alpha

Observation Measure	ELA	Mathematics	Combined ELA & Mathematics
Mastery Experiences	.85	.83	.65
Vicarious Experiences	.94	.71	.78
Verbal Persuasion	.83	-.14	.64
Physiological & Affective States	.85	.67	.81
Combined Subscales	.96	.84	.92

As explained in Chapter 3, the data collection period using this iteration of the tool corresponded with the Acuity C administration window. For this reason, the quantitative data collected for each indicator compared through inferential correlational analysis to the corresponding February/March survey administration results. In the following sections, I share the classroom level quantitative results of the correlation analysis between students' Acuity C outcomes in ELA and mathematics, the observation indicator data, and the survey data of the participants in those classes.

Mastery experiences. As explained previously in this chapter, I used Bandura's (1997) description of enactive mastery experiences (ME) to derive four success indicators during the coding process that I then used to collect quantitative data. Specifically, I attempted to capture the extent to which students: 1. *Experienced success*, 2. *Thought about / engaged with rigorous, grade-appropriate content (did the cognitive work)* 3. *Completed tasks that utilized prior knowledge or strategies to persevere when challenges arose*, and 4. *Appraised and monitored their progress*. Collectively, the ME subscale correlated to student self-efficacy levels in ELA ($r = .710, n = 10, p = .021$) and Acuity C mathematics outcomes ($r = .760, n = 10, p = .011$). Analysis revealed statistically significant relationships between self-efficacy in ELA and the *experienced success* ($r = .742, n = 10, p = .014$) and *monitor progress* ($r = .650, n = 10, p = .042$)

indicators, whereas Acuity C outcomes in ELA only correlated to the *monitor progress* indicator ($r = .719, n = 10, p = .019$). Mathematics achievement as well as self-efficacy levels displayed no significant correlations with any indicator in the ME subscale.

Vicarious experiences. As discussed previously in this chapter, Bandura (1997) presented research indicating students need to understand their own capabilities through self and peer appraisal, and learn to self-diagnose strengths and areas for improvement based on peer models to build self-efficacy. Therefore, I combined the four coded indicators to a single vicarious experiences (VE) subscale category. 1. *Used peer models for guidance*, 2. *Used peer models for self-evaluation or comparison*, 3. *Respectfully shared their thinking and differing opinions*, and 4. *Interacted with peers to support or complete work*. Analyses revealed a singular statistically significant relationship between *using peer models for guidance* and Acuity C achievement levels in ELA ($r = .636, n = 10, p = .048$). All other indicators within this subscale did not demonstrate correlations with self-efficacy or Acuity C achievement in ELA and mathematics. As a subscale, VE was not correlated to either self-efficacy or Acuity C outcomes in ELA or Mathematics.

Verbal persuasion. Verbal persuasion in the form of feedback and goal setting based on the quality of the effort, task, and learning relative to mastery rather than quantity or pace of progress was a documented source of self-efficacy (Bandura, 1997). Thus, I used my tool to collect evidence on four indicators of verbal persuasion (VP): 1. *Critiqued work and gave feedback toward the goal*, 2. *Assessed academic work*, 3. *Assessed work behaviors*, and 4. *Celebrated or offered encouragement for academic accomplishments and/or hard work*. Analysis of VP as subscale category revealed a statistically significant relationship with ELA self-efficacy levels ($r = .672, n = 10, p = .033$). Within this subscale, the *assess work behaviors*

($r = .695$, $n = 10$, $p = .026$) and *celebrate accomplishments* ($r = .733$, $n = 10$, $p = .016$) indicators also correlated with students' ELA self-efficacy levels. Analysis revealed a correlation between the *celebrate accomplishments* indicator and students' Acuity C outcomes in ELA ($r = .677$, $n = 10$, $p = .031$). There was a statistically significant correlation between self-efficacy in mathematics and the *critique work* indicator ($r = .660$, $n = 10$, $p = .038$), but no relationship between student outcomes on Acuity C and any of the indicators for the VP subscale ($p > .05$).

Physiological and affective states. As discussed in Chapter 2, people with self-efficacy tend to use struggle to learn how to handle their emotions related to unsuccessful experiences, and eventually view and manage them as opportunities (Bandura, 1997). It is a person's interpretation and integration of this feedback that determines the impact on self-efficacy (Bandura, 1997; Goddard, 2001). I used my tool to collect evidence on three indicators of students' physiological and affective states (PAS): 1. *Regulated emotion when it came to unsuccessful experiences*, 2. *Regulated/managed their behaviors*, and 3. *Managed their time effectively / were productive*. There was a statistically significant correlation between mathematics outcomes and the PAS subscale ($r = .653$, $n = 10$, $p = .041$), but there was not a relationship between the PAS subscale and ELA outcomes or self-efficacy in either subject. Student achievement in ELA on Acuity C did present a statistically significant relationship with the *manage time* indicator ($r = .659$, $n = 10$, $p = .038$), and achievement in mathematics correlated with both the *manage time* ($r = .821$, $n = 10$, $p = .004$) and *regulate behaviors* ($r = .652$, $n = 10$, $p = .041$) indicators.

Overall Question 4: What are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?

Data for this question were not directly collected during this study. Rather, syntheses of the data contribute to the overall understanding of this complex question. Therefore, question four is discussed in Chapter 5.

Summary

In summary, I conducted my concurrent triangulation mixed methods study in two phases. Phase I involved surveying 114 students and 10 teachers in 10 classrooms in grades three through five in two high poverty, urban schools. The quantitative survey results from the combined CORE District MESH and PALS surveys were used to measure the relationship between student self-efficacy levels and their perceptions of their teachers' effectiveness, and the teachers' perceptions of their students. While I found a statistically significant relationship ($p < .05$) between students' self-efficacy levels and their achievement in both administrations and both subjects, there was no such relationship with students' perceptions of their teachers' effectiveness. Teachers' perceptions of their students were only significant ($p < .05$) in regard to females.

Phase II was an instrument-design, sequential exploratory model involving observations, informal discussions with students, and document reviews. The purpose of this phase was to capture and refine lists of agentic and self-efficacious student behaviors to develop iterations of an observation tool for future revision and larger-scale validation. I found several statistically significant factors supporting my CASE theory and described them utilizing vignettes and pictures from the classroom. However, my observations were limited and warrant addition observations to confirm the results. Chapter 5 contains further discussions of these findings.

CHAPTER 5

DISCUSSION

This chapter contains the discussion of the results of my concurrent triangulation study of students' agentic and self-efficacious behaviors as an indicator of teacher effectiveness. This chapter begins with an explanation of the purpose and presentation of the research questions central to my study. I then summarize the major findings presented in Chapter 4 and discuss my research questions in the context of the findings of the literature. Finally, I explore the implications of this work and conclude with my recommendations for further research.

Purpose of the Study

The purpose of this concurrent triangulation mixed methods study (Creswell, 2014) was to identify the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in high poverty classrooms. In the first quantitative phase of the study, the research questions explored the relationship between student achievement, perceived student agency and academic efficacy in classrooms, and the students' levels of perceived teacher effectiveness. The independent variables in this phase were demographics, perceived self-efficacy, self-management, and social awareness, student perception of teacher, and teacher perception of student as measured by student and teacher surveys. The dependent variables were academic achievement in ELA and mathematics as measured by Acuity (McGraw Hill) assessments. The purpose of the concurrent second, qualitative phase of the study was to identify characteristics of classrooms and practices of teachers that establish agentic behaviors and academic self-efficacy beliefs of students.

Research Questions

The research questions this study responded to were:

Quantitative Phase I:

1. What are the relationships among students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness?

Qualitative Phase II:

2. What are the observable behaviors and norms of students in classrooms with varying levels of cultures of agency and self-efficacy (CASE)?
3. What relationship do students' agentic and efficacious classroom behaviors have with their perceived self-efficacy levels and academic achievement?

Overall:

4. What are the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms?

Review of the Research Methods

Within one school system and two campuses, I identified 10 high poverty classrooms ranging from grades three to five to conduct a concurrent multi-phase, mixed methods study. In the first quantitative phase of the study, I measured students' perceived agency and academic efficacy via three competencies on the CORE District survey (Krachman et al., 2016). Students' levels of perceived teacher effectiveness were measured by the Teacher Mastery Goal and Academic Press subscale portions of the PALS survey (Midgley et al., 2000). Upon completion of both collection phases, I conducted student-level analyses to answer the first research

question. Specifically, I used T-Tests to compare students' Acuity achievement descriptive data on the three ELA and mathematics tests (Acuity A, B, and C) with the discrete demographic indicators. I also analyzed the correlation between the survey results and student achievement and used regression to determine predictors. For the concurrent second qualitative phase of the study I collected data on the observable agentic behaviors of students in sample classrooms using coding to develop an observation tool, and described the results in Chapter 4 through vignettes. I used these Phase II data to triangulate correlations within Phase I data using classroom-level analyses to explore quantitatively the relationship between the characteristics of classrooms and practices of teachers that establish agentic behaviors and academic and self-efficacy beliefs of students.

Findings Related to the Literature

This section contains the discussion of my findings in relation to the conclusions of other researchers. Following a brief discussion related to the demographic findings, the remainder of the section is organized by question. My final research question is discussed in greater detail and synthesizes my findings for questions one through three as an exploration into the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms.

Demographics

There were only three demographic variables in my study that presented statistically significant correlations with Acuity outcomes: gender, looped status, and special education enrollment. My study found females scored higher than males in the social awareness indicator on the survey in February/March. This was consistent with the findings of Krachman et al.

(2016) in their analysis of the CORE district survey results from almost half a million students in grades 4-12 in California. However, contrary to the findings of Pina-Neves et al. (2013) and Krachman et al. (2016) who found males had higher self-efficacy levels than females in their separate studies of Portuguese high school students and the CORE districts in California respectively, my study found no significant correlation between students' gender and their self-efficacy levels.

A study conducted by Gresham and Elliot (2001) found students in grades three, four, and five with mild disabilities self reported lower self-efficacy levels than their non-disabled peers. A study by Baird, Scott, Dearing, and Hamill (2009) also found students with learning disabilities had lower academic self-efficacy levels than their non-disabled peers and were more likely to view their knowledge and intelligence as a fixed measure. However, my study found no significant differences in self-efficacy levels based on special education enrollment. And, unlike the findings several studies including those from Judge and Watson (2011) and Schulte and Stevens (2015), my study did not find special education enrollment correlated with lower outcomes in mathematics, only ELA.

I found looped status correlated with Acuity outcomes on all administrations in both subjects. Specifically, students who looped with their teacher scored higher on all Acuity assessments than those in classrooms that did not loop. These results were consistent with findings of Cistone and Shneyderman (2004), Hill and Jones (2018), and Lincoln (1997). Cistone and Shneyderman (2004) noted the importance of looping in building relationships between teachers, students, and their families and the correlation between these strong relationships and increases in student learning. Hill and Jones (2018) found effects were largest for minority students who looped. While ethnicity was not a significant demographic in my

study in relation to Acuity outcomes, 82% of my sample population were minorities. While looped status was not significantly correlated with self-efficacy levels, I could not find any literature to date studying the relationship between looping and self-efficacy levels.

Quantitative Phase: Research Question One

My first research question addressed the relationships between students' academic achievement, their level of academic agency and self-efficacy, and their perceptions of their teachers' effectiveness. I found a positive correlation between students' perceived self-efficacy and their achievement in both ELA and mathematics, which is consistent with literature (Farrington et al., 2012; Foster, 2015; Mega et al., 2014; Multon et al., 1991; Pajares & Kranzler, 1995; Schunk, 1989; West, 2014; Zahodne et al., 2015; Zimmerman, 1999). I also found the self-efficacy indicator to be the best predictor of achievement in both subjects as well. Conversely, Krachman et al. (2016) found the self-management indicator on the CORE Districts survey to be the best predictor of ELA achievement in elementary and middle school. Furthermore, self-efficacy explained 15-17% of the variance in outcomes in ELA and mathematics in my study when controlling for looped status and SpEd enrollment. This percentage is slightly higher than the 14% found by Multon et al. (1991) in their meta-analysis of 40 self-efficacy studies.

Although not explicitly tied to this research question, I administered the teacher perception of students (TPS) portion of the CORE Districts' survey in part to confirm students' perceptions of their own self-management (SM) and social awareness (SA). Similar to the research conducted by Krachman et al., (2016) and West (2016), my study found a correlation between TPS, SM, self-efficacy, and student achievement. Unlike these studies, I did not find a correlation with SA. Similar to my study, Rosenholtz and Wilson (1980) as well as Rosenholtz

and Simpson (1984) also found that teachers' assessments of their students' abilities demonstrated a strong correlation with students' own assessments of their personal academic abilities.

Additionally, I found no prior studies that had paired the CORE Districts survey (Krachman et al., 2016) and PALS survey (Midgley et al., 2000) to establish correlations between students' their perceptions of their teachers and their self-efficacy and achievement. While my study explored that relationship, none was found.

Qualitative Phase: Research Question Two

The intent of my second research question was to capture the observable behaviors and norms of students in classrooms with varying levels of agency and self-efficacy. While I presented the results for this question in Chapter 4 as a methodical accounting of each observed indicator through the use of examples and vignettes, the following section is a synthesis and reflection on these findings according to classrooms above and below the mean for student self-efficacy. Due to the inductive nature of this process, it was difficult to connect my findings to the literature exclusively in the context of observable behaviors and without literature that correlated them to self-efficacy and academic achievement outcomes. However, a separate section exploring my findings in more detail as they relate to question three and the correlations in my study follows this section.

According to Bandura (1997), teachers could help students control two areas related to self-efficacy, cognition and behavior. He also noted that there is an inherent relationship between people, their environments, and the outcomes that are derived from their interactions (Bandura, 2006). During my rounds of observations I focused on documenting the behaviors

and cognition of the students rather than those of the teacher, and capturing evidence of the cultivated learning environments.

Additionally, during my Phase II research, I concluded students' demonstration of these behaviors independent of or co-dependent with the teacher was an important variable to capture. Similar to Dorman's (2001) study that found a correlation between the "extent to which students are invited to share with the teacher control of the learning environment" (p. 248) and their academic self-efficacy levels, I found students in higher efficacious classrooms were given more opportunities for self-governance, autonomy, and self-regulation than those with levels below the mean. This is consistent with Bandura's (1997) and Lüftenegger et al.'s (2016) research that students feel a higher sense of self-efficacy when they are accorded a level of license to explore, manipulate, and influence their learning environments. Specifically, Lüftenegger et al. (2016) found students "are more interested, learning goal oriented and show higher self-efficacy when they perceive themselves as more self-determined and autonomous in classroom learning activities" (p. 34).

Higher efficacious classrooms. There were several common environmental features and self-efficacious student behaviors I found in classrooms with students reporting self-efficacy levels higher than the mean. My findings were mostly consistent with those of other researchers and I noted the supporting literature after each. I discuss disconfirming findings at the conclusion of these lists. In higher efficacious classrooms:

Mastery Experiences -

- Students engaged in work and conversations with their teachers and peers that emphasized learning and learning processes (Ericsson & Charness, 1994; Lodewky & Wynn, 2005; Schunk 1989).

- Students typically worked on tasks that were at or above grade level (Bandura, 1997; Pajares, 2008; Putney & Broughton, 2011).
- Students relied on prior knowledge to solve problems or conceived of strategies to help them in the process. They typically had ready access to learning tools in the classroom (such as manipulatives, computers, rulers, dictionaries, etc.) and used them (Bandura et al., 2003; Pintrich & De Groot, 1990; Ryan et al., 1998; Schunk, 1989; Zimmerman, 2002).
- Students self-monitored and reflected on their progress towards learning and mastering the content, and structures, norms, and routines for them to do so were consistently present (Schunk & Rice, 1984; Zimmerman, 2002).

Vicarious Experiences –

- Students had frequent opportunities for peer learning and used their peers to develop their own strategy and content knowledge (Bandura, 1997; Hushman, & Marley, 2015).
- Students used their peers' work as exemplars and anchors to inform their work quality and learning, and/or to correct misconceptions (Pajares, 1996).
- Students appeared to have strong relationships with each other and had routines and norms for discourse and disagreement (Patrick et al., 1997; Schwartz et al., 2005; St Clair-Thompson et al., 2015; Tulis et al., 2018).

Verbal Persuasion –

- Students had time and structures for giving feedback to one another based on goals and/or quality criteria (Bandura, 1997; Schunk, 1987).

- Students had an understanding of their academic and behavioral strengths and weaknesses, and had structures and supports for setting their own goals and reflecting on their progress (Bandura, 2012; Pintrich & De Groot, 1990; Putney & Broughton, 2011; Schunk, 1987; Zimmerman, 2011).
- Students engaged in celebrations of their peers' work and behaviors, including academic risk-taking, perseverance, and success (Bandura, 1997; Zimmerman, 2011).

Physiological and Affective States –

- Students had strategies for self-regulating their emotions and behaviors, and the freedom to pursue self-calming measures as needed (Graziano, Reavis, Keane, Calkins, 2007; Zimmerman, 2002).
- Students received feedback and help from the teacher and/or their peers to strategize, set goals, and plan to overcome barriers to success (Bandura, 1997; Bandura & Schunk, 1981; Schunk, 1987; Zimmerman & Kitsantas, 1999).
- Students and teachers had structures to manage their time well. Students had clear norms and routines for what they should do if they finished with a task early, and generally student learning and understanding of the academic content guided the pace (Bouffard-Bouchard et al., 1991; Multon et al., 1991; Zimmerman, 2002, 2011).

Lower efficacious classrooms. Conversely, there were several common environmental features and self-efficacious student behaviors I found in classrooms with students reporting self-

efficacy levels lower than the mean. They are listed below and are organized by source category.

The supporting literature is noted after each finding. In lower efficacious classrooms:

Mastery Experiences -

- Student and teacher conversations and interactions emphasized correct answers. (Discourse, modeling, teacher prompting were focused on the correct answers for the task rather than how the student would arrive there.) (Ericsson & Charness, 1994; Lodewyck & Wynn, 2005; Schunk 1989; Schunk & Rice, 1987).
- Students worked on tasks that were only partially aligned or were not aligned to grade level standards (Bandura, 1997).
- Students struggled to problem solve using tools or prior knowledge. They tended to wait to be told/shown what to do or how to do it, frequently quit working, and/or did not utilize learning tools in the classroom (Bandura & Schunk, 1981; Pajares, 1996; Pajares & Johnson, 1994).
- Students did not have norms or routines for self-appraisal or to monitor their progress, or they were generally generic (such as being prompted to give a thumbs up if they were finished with a task or ready to move on) (Schunk, 1996; Schunk & Rice, 1989).

Vicarious Experiences –

- Students had limited opportunities to interact with their peers or utilized them as sources for correct answers. Teacher-talk or didactic interactions tended to dominate these classrooms (Hushman, & Marley, 2015; Schunk, 1981).

- Structures for peer comparisons were absent or they tended to be devoid of specific context (appraisals were generic in nature and tended to focus on students doing better or worse than their peers) (Schunk & Zimmerman, 1998)
- Students appeared to have modest relationships with each other. Structures or norms for discourse and disagreement tended to flow through the teacher rather than peer to peer (Hushman, & Marley, 2015; Patrick et al., 1997; Schwartz et al., 2005; St Clair-Thompson et al., 2015; Tulis et al., 2018).

Verbal Persuasion –

- Students lacked time and/or structures for providing feedback to one another (Bandura, 1997; Schunk, 1987).
- Teachers tended to identify the academic strengths and weaknesses for the student, or the students were generally unaware of what they were. Feedback tended toward behavior or effort toward the task (Bouffard-Bouchard et al., 1991; Schunk, 1984, 1989; Zimmerman & Kitsantas, 1999).
- Students usually did not actively engage in academic or behavior celebrations. Typically, celebrations were the responsibility of the teacher and tended to focus on reinforcing compliance, effort behaviors, and correct answers (Bandura, 1997; Bouffard-Bouchard et al., 1991).

Physiological and Affective States –

- Students struggled with self-regulating their emotions and/or academic behaviors (Schunk & Zimmerman, 1998).

- The teacher frequently intervened to correct students' negative behaviors or remove/lessen sources of agitation and frustration/barriers to learning for students (Bandura, 1997).
- Students and/or teachers struggled with time management. If class pacing lagged, students typically did not have routines or norms for what they should do if they finished a task early, or the expectation was for them to sit quietly. When pacing was too brisk, the emphasis tended to be placed on task completion rather than accuracy or understanding (Bouffard-Bouchard et al., 1991; Multon et al., 1991; Zimmerman, 2002, 2011).

Disconfirmation of My Findings. One major finding in my study ran counter to other researchers' findings. Zimmerman and Kitsantas (1999) found students who focused on getting better outcomes in a writing exercise out performed and had higher self-efficacy levels than students who focused on improving their processes. Conversely, I found classrooms with higher self-efficacy levels were more process-focused than those classrooms with lower self-efficacy levels, which tended to have teachers and students more focused on outcomes and correct answers. This may be due to what Schunk (1989) and Pajares (1996) described as a difference between self-efficacy for performance versus self-efficacy for learning. Lodewky and Wynn (2005) concluded activities that improve self-efficacy for learning had greater alignment to the task level, which may account for my observations during Phase II. My qualitative findings were therefore more consistent with Lodewky and Wynn (2005) Schunk (1989), Schunk and Rice (1991), and Pajares (1996) than Zimmerman and Kitsantas (1999). Specifically, the difference between my findings and Zimmerman and Kitsantas (1999) may in part be attributed to the general consistency within the high efficacious classrooms of allowing students choice in

their processes while emphasizing the “how and why” behind them as opposed to the explicit refinement or mastery of a particular strategy.

Qualitative Phase: Research Question Three

My third research question focused on whether there were relationships between students’ agentic and efficacious behaviors in classrooms and their perceived self-efficacy levels and academic achievement. During my analysis I found all of the observation source subscales to have internal consistency and reliability as measured by Cronbach’s alpha with the exception of *verbal persuasion* in mathematics. I also found relationships within three of the subscales and between seven of the indicators with students’ self-efficacy levels and academic outcomes in ELA and/or mathematics. I was hesitant to present findings that disconfirmed the literature due to the limited scope and small sample size of my study. I do not believe that the findings that lacked statistical significance in my study in any way disprove or call into question other researchers’ work. Bandura (2012) emphasized that studies that disconfirm established research tend to analyze self-efficacy from the setting, domain, or general levels. While my observations were conducted to collect data at the task level, the surveys collected efficacy levels at the domain level. My analysis compared data collected at varying levels of specificity and thus may also have contributed to inconsistencies in my findings relative to the literature. My findings in the context of the literature are presented in the following paragraphs.

Mastery Experiences. The *mastery experiences* subscale in my study correlated to self-efficacy levels in ELA and students outcomes in mathematics. As noted by several researchers, mastery experiences help build confidence and perceptions of skillfulness, which then effect efficacy levels and approaches to future learning (Bernacki et al., 2015; Usher & Pajares, 2006; Zimmerman, 2011). Through their study of 395 third graders in 21 classrooms, Joët, Usher, and

Bressoux (2011) found that mastery experiences were predictive of achievement in mathematics, and although they worked with older students, Bernacki et al. (2015) also found a correlation between students' mastery experiences in mathematics tasks, self-efficacy levels and their outcomes. Both of these studies helped confirm my results correlating this subscale to mathematics achievement. Additionally, Pajares, Johnson and Usher (2007) found a correlation between mastery experiences and student self-efficacy levels in writing, similar to my findings in ELA.

The *experience success* indicator within this subscale correlated with students' self-efficacy levels in ELA. There was no correlation to mathematics self-efficacy or achievement. Bandura (1997) described the self-perpetuating process success plays in building a students' sense of self-efficacy; that if tasks were too easy they had little contribution but if they were deemed challenging they could raise or lower a students sense of self-efficacy depending on the outcome. Challenging tasks that were met with occasional failure could improve self-efficacy levels if the student viewed themselves as successfully overcoming and mastering the task. Reciprocally, how a student approached a problem and the extent to which s/he persisted was dependent on his/her self-efficacy levels (Bandura, 2012). Similar to my findings, Schunk (1981) also found that when students have the independence to apply their skills in new situations unaided by a teacher, they feel more successful than when they have that additional support. Limitations in the number of observations conducted may account for the lack of similar correlations in mathematics in my study.

The *appraise and monitor progress* indicator also correlated with students' self-efficacy levels and achievement in ELA. For this indicator, I looked for evidence that the students assessed their progress towards their success on the task, targets, and/or behavior objectives. My

findings were similar to Schunk (1990) who found self-efficacy levels were correlated to students setting goals, perceiving progress towards the goals, and students feeling capable of achieving those goals. Schunk (1985) also noted the practice of engaging in goal setting conferences as particularly effective for improving self-efficacy levels because they contribute to students' sense of control over their outcomes. This was consistent with my findings in ELA and many of my observations in classrooms. Limitations in the number of observations may explain the lack of correlation in mathematics.

Vicarious Experiences. Analysis revealed *vicarious experiences* as a subscale had no statistically significant relationship with self-efficacy or student achievement in either ELA or mathematics. This is inconsistent with much of the literature (Bandura, 1997; Pajares, 2003). However, in their study of elementary, middle, and high school students and their self-efficacy levels in writing, Pajares et al. (2007) did not find a correlation between self-efficacy and their measures of vicarious experiences, which was consistent with my findings. The only indicator within this subscale to present a correlation was *used peer models for guidance*. This indicator had a statistically significant relationship with ELA achievement exclusively. Similarly, in their study of self-efficacy levels of English Language Learners (ELL), Boisvert and Rao (2015) found students who used peer models for guidance accelerated their learning process and presented increased self-efficacy levels.

Verbal Persuasion. The *verbal persuasion* subscale as I measured it correlated only with ELA self-efficacy levels in my study, similar to the findings of Schunk (1984), Pajares (2003), and Pajares et al. (2007) who also found correlations between verbal persuasion and self-efficacy levels in ELA. Similar to my study, Matsui, Matsui, and Ohnishi (1990) found no correlation between verbal persuasion and mathematics self-efficacy levels in students. The

assessed work behaviors indicator within this subscale correlated with self-efficacy outcomes in ELA, which was consistent with the literature (Bandura, 2012; Pintrich & De Groot, 1990; Putney & Broughton, 2011; Schunk, 1987; Zimmerman, 2011). The *Celebrated or offered encouragement for academic accomplishments and/or hard work* indicator presented a statistically significant relationship with both self-efficacy levels and academic outcomes in ELA in my study, which was also consistent with the literature (Bandura, 1997; Zimmerman, 2011).

Physiological and Affective States. I found the *physiological and affective states* subscale correlated with mathematics outcomes exclusively. These findings in mathematics were consistent with Lopez and Lent's (1992) findings in their study of high school students and with Joët et al.'s (2011) findings. However, in their study of nearly 400 third grade students in France, Joët et al. (2011) found all four sources of self-efficacy, including their measures of the *physiological and affective states* source, influenced the participants' perceived self-efficacy levels as well as correlated to their achievement levels in both French and mathematics. This was inconsistent with my results, which found no correlation between PAS with self-efficacy levels in either subject or achievement in ELA.

Within the *physiological and affective states* subscale, the *regulated/managed their behaviors* indicator presented a statistically significant relationship with mathematics achievement exclusively. The lack of correlation with self-efficacy was inconsistent with the research that found regulating behaviors and emotions (Bandura, 2012; Putney & Broughton, 2011; Zimmerman, 2011), including anxiety (Pajares, 2003; Pintrich & De Groot, 1990), correlated with self-efficacy levels. Additionally, the *managed their time effectively / were productive* indicator in my study correlated to student achievement in both ELA and mathematics, which was supported by studies conducted mostly in high school and college

(Britton & Tesser, 1991; Lane & Lane, 2001) or meta-analysis (Multon et al., 1991). However, this indicator did not correlate to self-efficacy levels in either subject, which ran counter to the literature (Bandura, 1997; Bouffard-Bouchard et al., 1991; Cleary & Zimmerman, 2004; Schunk, 1989; Zimmerman 2002, 2011).

Research Question Four

My overarching research question explored the teacher and student practices or behaviors that support cultures of agency and self-efficacy (CASE) in classrooms. Rather than consider the teacher-centric technical and procedural components of the classroom that Martinez et al. (2016) described as typical of American systems, I attempted to follow the lead of the world's more progressive and some of the top-ranked education systems, namely Singapore and Japan, who practice more holistic, learner-centric self-reflection and evaluation. Namely, I attempted to describe the role of the teacher by studying and articulating the student agentic and self-efficacious behaviors s/he should attempt to cultivate in students. This student-centric approach is also more in line with what Ovando (2001) described in her research of learner-centered evaluation systems. Specifically, she found that when teachers received evaluative feedback and professional support based on data derived from the students' behaviors and outcomes rather than their own they had a greater awareness of their students' needs and more freedom to be nimble and responsive to those needs; the hallmarks of a strong and effective evaluation system (Gullickson & Howard, 2009).

Researchers have shown that learning environments established by teachers and the extent to which students were given opportunities to share control and self-regulate correlated with higher students' self-efficacy levels (Dorman, 2001; Yerdelen & Sungur, 2019). This paralleled the findings in my study. Furthermore, my qualitative research revealed students

shared a degree of interdependence in those classrooms with higher levels of self-efficacy. They had the freedom to converse and interact with each other about and to progress their learning. They had structures and processes in place for collective and self-reflection, assessment, and goal-setting. They were empowered to utilize the tools available in the classroom (including their peers) to meet their own individual needs, yet also contributed to the supports of the collective through peer feedback, collaboration, and celebrations. Students in these classrooms generally took more academic risks, had greater responsibilities for their own and collective work, and self-regulated their processes of learning and for controlling their emotions, responding positively or neutrally to failure and set-backs.

I found these behaviors in classrooms from a range of grades three through five and concluded these behaviors are not exclusive to the older students. Furthermore, looped classrooms were more likely to have students present these behaviors than classrooms that were not in their second year with a teacher. The impact of looping (discussed later in implications for practice section) on self-efficacy, while statistically nonsignificant, is still notable, especially considering the correlation found with achievement. My findings indicated students who looped with their teacher had higher self-efficacy and achievement levels than those who did not. This could be attributed to Bernacki et al. (2015) and Putney's (1996) assertions that self-efficacy is not a continuous, consistent construct but a complex and dynamic continuum informed by multiple variables and inputs that evolves over time.

These student behaviors and the commonalities in cultures of these classrooms have implications on role teachers play in building cultures of agency and self-efficacy (CASE) in their classrooms. The seminal piece of literature grounding this portion of my discussion was Putney and Broughton's (2011) longitudinal qualitative study of Ms. Fall's fifth grade classroom.

Very similar to their findings, I found classrooms with the greatest levels of perceived student self-efficacy had teachers who took on the role of community organizers (Bandura, 1997; Putney & Broughton, 2011). Specifically, those teachers provided space and developed structures to enable their students to interact constructively with peers and content in ways that nurtured and reinforced their sense of agency and self-efficacy not just as singular individuals, but as a system of collaborative agents for collective learning. Whereas I called this a culture of agency and self-efficacy (CASE), Putney and Broughton (2011) combine this construct with teacher efficacy and coined it collective classroom efficacy.

Limitations

My study had several limitations. First, my sample size was small. In some cases, classrooms only had five participants from whom I collected Phase I data. Additionally, I drew conclusions both in my Phase II collection window and the classroom-level analysis I conducted to help answer my third research question based on a sample size of 10 classrooms. Hox (2002) recommends at least 40-50 classrooms to effectively and reliably estimate group effects. Second, the case study schools had little demographic variance between each other. They were very homogeneous in the SES of their student populations. While this study may be beneficial for schools with significantly high populations of low SES students, it would need replication in schools with more diverse and larger sample groups. Third, there was no prior relationship established between the CORE District survey and the PALS survey. This would need further study to determine if my null results were typical. Fourth, the timing of the IRB process meant my study was truncated from the original timeline. The Phase II process was particularly short and data used to answer my third research question was mostly collected during a single observation in each subject area. Finally, the tool created during the last stage of Phase II needs

to be revised and tested for further validity and reliability across a larger and more diverse sample.

Implications and Recommendations for Practice

In review, a student's agency and efficacy beliefs are affected by their behavior, the environment, and other personal factors including biology, cognition, and affective events (Bandura, 1997). Teachers have enormous control and influence in their classrooms (Kane & Staiger, 2008; Wright, Horn, & Sanders, 1997). To overcome any negative influences in factors outside of direct control, a teacher in a high poverty, urban school must teach students how to regulate their behavior, manipulate their environments to be successful, and support development of positive personal factors. Many students from chronic poverty bring, often underappreciated, skills to the classroom, which contribute to their level of efficacy and agency outside of academia (Berliner, 2013; Socol, 2014). Many of these students are resilient, adapting and surviving circumstances and life events that would intimidate many adults. They take on adult roles and responsibilities at home and act as caregivers for younger siblings. They are resourceful and creative in overcoming situations, and have an ability to conceive of opportunities where others would see none (Mullainathan & Shafir, 2013; Socol, 2014). These traits may facilitate academic agency and efficacy if these students can transfer their resilience and adaptability to their academic pursuits (Socol, 2014).

It is important for teachers to foster learning environments and behaviors such as self-regulation and metacognition (Dorman, 2001; Yerdelen & Sungur, 2019) that help with transfer of student understanding and non-cognitive skills to academic contexts, where efficacy is cultivated and fostered (Farnham et al., 2015; Phan, 2013; Schunk & Rice, 1987; Tough, 2015). Findings from Schunk and Rice (1987) suggested teachers should not only create lessons and

activities that teach multiple strategies to approach problem solving and interpret feedback, but also explore explicitly why they are effective and under which circumstances. My research supported this finding. I also found teachers with higher levels of student self-efficacy had students who were able to recognize why a strategy was useful and when to apply it, and these teachers gave their students regular opportunities to choose when and how to deploy the strategies giving them a degree of control over their learning, similar to the findings of Schunk and Rice (1987). Teachers could then look for these student behaviors to help inform whether their classroom environment and learning tasks are contributing to the cultivation of efficacy in their classrooms. If teachers could explicitly establish learning environments, teach, and support the means that allow students to accomplish tasks and act in ways that cultivate success, students could feel more in control of their own development and increase agency outcomes and efficacy beliefs (Bandura, 1997; Schunk & Rice, 1987).

Recommendations for Teachers

Based on my findings, I encourage teachers to consider norming classroom conversations and students' peer-to-peer interactions to help ensure are academically productive and emphasize learning and learning processes. I also recommend teachers ensure students work on tasks that are at or above grade level, and utilize tasks and instructional strategies that require students rely on prior knowledge conceive of strategies to help them problem solve. This could involve asking students to complete tasks that have them grapple with content slightly beyond their current level of mastery of a topic so they have to utilize prior knowledge rather than a teacher-modeled strategy to work their way to a solution. Additionally, making learning tools, including peers, readily available in classrooms and teaching students how to use them effectively could also benefit students.

I recommend teachers intentionally select strategies to support students' metacognition to help them self-monitor their learning and reflect on their progress towards mastering content. They should create structures, norms, and routines for students to engage in these activities regularly. As a part of this process, teachers should ensure students receive feedback and help from both them and their peers, and that students have strategies, time, and structures to strategize, set their own goals, craft plans to overcome barriers to success, and reflect on their progress. This feedback should also help students understand their academic and behavioral strengths and weaknesses. Teachers should have structures to manage their time well and teach and expect students to manage theirs as well. This could involve providing students with clear norms and routines for what they should do if they finished with a task early, and using student learning and understanding of the academic content to guide the pace of instruction. Teaching students strategies for self-regulating their emotions and behaviors, and providing opportunities for them to pursue self-calming measures as needed would also benefit students and help build a culture of self-efficacy in classrooms.

Furthermore, I urge teachers to support and encourage their students to have strong relationships with each other and develop norms and routines for academic discourse and disagreement. As a part of this process, provide students with time and structures for giving feedback to one another based on goals and/or quality criteria. Teachers should also create frequent opportunities for students to engage in peer learning and encourage them to use their peers to develop their own strategy and content knowledge. I also recommend teachers make it a regular habit for students to deconstruct and use their peers' work as exemplars and anchors to inform their work quality and learning, and/or to correct their misconceptions. Teachers should

empower students to engage in celebrations of their peers' work and behaviors, including academic risk-taking, perseverance, and success.

Recommendations for Leaders

Based on my findings, there are a few recommendations I would encourage school leaders take to help teachers and students get better academic and social outcomes. From a structure standpoint, I encourage school leaders consider orienting teacher support rubrics, tools, and evaluation measures toward measuring the degree to which students and not teachers demonstrate the ideals of the classroom environment, moving from teacher procedures to student competencies. I also encourage leaders to consider looping students with their teacher as a relatively low cost measure for improving student outcomes and self-efficacy levels.

There are also a few recommendations I encourage school leaders, teacher leaders, and coaches consider in regards to the support they provide teachers. Consider collecting data and including measures of student self-efficacy in the feedback provided to teachers to support their improvement efforts. I also recommend that during classroom walk-throughs, observations, and evaluations, leaders and coaches collect evidence on the classroom cultures based on student behaviors to provide coaching and feedback to teachers on their roles as community organizers (Bandura, 1997) in their classrooms. Finally, given the consistent correlations found between student self-efficacy and their achievement outcomes, I encourage leaders consider the inclusion of student self-efficacy measures as an indicator of teachers' effectiveness in the classroom.

Implications for Theory

My CASE conceptual framework maintains that between Bandura's (1997) measurement of the individual self-efficacy of students, and the collective teacher efficacy of the organization,

there exists a measurable level of agency and self-efficacy cultivated by the teacher and students in the classroom environment. Furthermore, the evaluation of the agentic classroom culture can be utilized to measure a teacher's effectiveness, similar to the way 'respectful culture' is currently utilized by several prominent evaluation systems as in indicator of effectiveness (Danielson, 2013; National Institute for Excellence in Teaching, 1999). This agentic and self-efficacious culture includes norms, beliefs, behaviors, and language that build and indicate the presence of student agency and academic efficacy as presented in my study. I believe there is a gap in the research at the classroom level, particularly in qualitative research. While I recognize this presents a challenge from both a research on child participants and a scope/scale perspective, I believe it is necessary to better understand the dynamic between student self-efficacy levels, their presenting behaviors, and the teacher responsibilities for cultivating those behaviors.

Recommendations for Further Research

This study was limited by both time and sample size. The Phase II process was particularly short and data used to answer my third research question was mostly collected during a single observation in each subject area within 10 classrooms. As previously discussed, Hox (2002) recommends data collection in at least 40-50 classrooms to effectively and reliably estimate group effects. I recommend expansion of the sample group to one of this size. Replication of Strong's (2011) methods utilizing video may be an effective way manage the larger sample. Inclusion of schools with more diverse and larger sample groups may also be helpful to better understand greater demographic implications. Although the collection tool I developed in Phase II of this study presented internal consistency of all but one subscale in mathematics, further refinement of the collection tool is necessary. Specifically, I recommend further coding and simplification of the indicators, and further testing for validity and reliability

across a larger and more diverse sample. One possibility for further coding and simplification includes organizing the tool by the four properties described by Bandura (2006), intentionality, forethought, self-reactiveness, and self-reflectiveness, rather than by source.

I also recommend future research involve a longitudinal study of looped classrooms to better capture the process involved in establishing CASE. Putney and Broughton's (2011) longitudinal study was of a single teacher with varying student groups, and included measuring the dynamic between teachers' collective efficacy and individual study self-efficacy as it was manifested in collective classroom efficacy. My study did not take into account teacher collective efficacy levels. This could be a measure included in the future to help better understand the dynamic between these measures and constructs. A longitudinal study of looped classrooms for evidence of CASE paired with data collection on teacher efficacy could be helpful for the field.

REFERENCES

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25(1), 95–135.
- Afolabi, K.A. (2010). Relationship of self-efficacy beliefs of urban public school students to performance on a high-stakes mathematics test. University of Massachusetts. Dissertations. Paper 262. Retrieved from http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1260&context=open_access_dissertations
- Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research*, 3, 101–134.
- Allen, D., & Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE-Life Sciences Education*, 5(3), 197-203.
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034–1037.
- Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009). *The schools teachers leave: Teacher mobility in Chicago Public Schools*. Chicago: Consortium on Chicago School Research.
- Arter, J. A., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, Calif: Corwin Press.
- Baird, G. L., Scott, W. D., Dearing, E., & Hamill, S. K. (2009). Cognitive self-regulation in youth with and without learning disabilities: Academic self-efficacy, theories of intelligence, learning vs. performance goal preferences, and effort attributions. *Journal of Social and Clinical Psychology*, 28(7), 881-908.

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, 44(9), 1175-1184.
- Bandura, A. (1997). *Self-Efficacy: The exercise of control*. New York: Freeman
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1(2), 164-180.
- Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, 38(1): 9-44.
- Bandura, A., Caprara, G. V., Barbaranelli, C., Gerbino, M., & Pastorelli, C. (2003). Role of affective self-regulatory efficacy in diverse spheres of psychosocial functioning. *Child development*, 74(3), 769-782.
- Bandura, A. & Schunk, D. H. (1981). Cultivating competency, self-efficacy and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41, 586-598.
- Berliner, D. C, (2013). Inequality, poverty, and the socialization of America's youth for the responsibilities of citizenship. *Theory into Practice*, 52(3), 203-209.
- Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning. *Metacognition and Learning*, 10(1), 99-117.
- Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *Elementary School Journal*, 111(1), 7-34.

- Bill & Melinda Gates Foundation. (2010). *Working with teachers to develop fair and reliable measures of effective teaching*. Retrieved from <https://docs.gatesfoundation.org/Documents/met-framing-paper.pdf>
- Bill and Melinda Gates Foundation. (2011). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Bellevue, WA: Author. Retrieved from www.gatesfoundation.org/collegeready-education/Documents/preliminary-findings-research-paper.pdf
- Bloomberge, L. & Volpe, M. (2016). *Completing your qualitative dissertation: A road map from beginning to end*. Thousand Oaks, CA: Sage.
- Boisvert, P. & Rao, K. (2015). Video self-modeling for English Language Learners. *TESOL Journal*, 6(1). 36-58.
- Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, 93(1), 23-34.
- Bong, M., & Clark, R. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, 34(3), 139–154
- Bouffard-Bouchard, T., Parent, S., & Larivée, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development*, 14(2), 154-164.
- Briggs, D. & Domingue, B. (2011). *Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District teachers by the Los Angeles Times*. Boulder, CO: National Education Policy Center.

- Britton, B.K., & Tesser, A. (1991). Effects of time-management practices on college grades. *Journal of Educational Psychology*, 83, 405-41.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.
- Carnevale, A. P., & Strohl, J. (2010). How increasing college access is increasing inequality, and what to do about it. In R. D. Kahlenberg (Ed.), *Rewarding strivers: Helping low-income students succeed in college* (pp. 71–183). New York, NY: Century Foundation Press.
- Cassidy, S. (2015). Resilience building in students: the role of academic self-efficacy. *Frontiers in Psychology*, 6.
- Chapman, M., Skinner, E. A., & Baltes P. B. (1990). Interpreting correlations between children's perceived control and cognitive performance: Control, agency, or means-ends beliefs? *Developmental Psychology*, 26(2): 246-253.
- Chirichello, M., & Chirichello, C. (2001). A standing ovation for looping: The critics respond. *Childhood Education*, 78(1), 2.
- Cistone, P., & Shneyderman, A. (2004). Looping: An empirical evaluation. *International Journal of Educational Policy, Research, and Practice: Reconceptualizing Childhood Studies*, 5(1), 47-61.
- Cleary, T. J., & Callan, G. L. (2014). Student self-regulated learning in an urban high school: Predictive validity and relations between teacher ratings and student self-reports. *Journal of Psychoeducational Assessment*, 32(4), 295-305.
- Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools*, 41(5), 537-550.

- Clements, A., & Selznick, B. (2007). *Lunch money*. New York: Atheneum Books for Young Readers.
- Coleman, J. (1966.) Equality of Educational Opportunity. *National Center for Educational Statistics*. Retrieved from <http://files.eric.ed.gov/fulltext/ED012275.pdf>
- Collins, J L. (1985). Self-efficacy and ability in achievement behavior. (Order No. 8506172). Available from ProQuest Dissertations & Theses A&I. (303396520).
- Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., McLean, L., & Morrison, F. J. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. *Journal of educational psychology*, 106(3), 762.
- Corcoran, K. E., Pettinicchio, D., & Young, J. T. N. (2011). The context of control: a cross-national investigation of the link between political institutions, efficacy, and collective action. *British Journal of Social Psychology*, 50(4), 575-605.
- CORE Districts (2013). About CORE. Retrieved February 18, 2017 from <http://coredistricts.org/why-is-core-needed/>
- CORE Districts (2016a). Social-emotional skills. Retrieved from <https://s3.amazonaws.com/uploads.hipchat.com/392387/2399075/U3GNMByhpldya9K/S-E-CC-Domain-Social-Emotional-Skills-updated-3.29.16.pdf>.
- CORE Districts (2016b). The school-quality improvement index & CORE data collaborative. Retrieved from <http://coredistricts.org/wp-content/uploads/2016/01/CORE-Data-Collaborative-v3-1-21-16.pdf>.

- Credé, M., Tynan, M. C., & Harms, P. D. (2016). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*. Retrieved from <http://dx.doi.org/10.1037/pspp0000102>.
- Creswell, J.W. (1999). Mixed-methods research: Introduction and application. In *Handbook of educational policy* (pp. 455-472). Cambridge, MA: Academic Press.
- Creswell, J.W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, N.J.: Merrill.
- Creswell, J.W. (2014). Research design: *Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, California: SAGE Publications.
- Creswell, J.W., Plano Clark, V.L. (2017). *Designing and conducting mixed methods research*. Thousand Oaks, California: SAGE Publications.
- Creswell, J.W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, CA: Sage.
- Danielson, C. (2013). 2013 Framework for teaching evaluation instrument. Retrieved from <https://www.danielsongroup.org/download/?download=448>.
- Danielson C. & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher* 44(2), 132-137.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kapan*, 93(6), 8-15.

- Darling-Hammond, L., Bae, S., Cook-Harvey, C. M., Lam, L., Mercer, C., Podolsky, A. & Leisy Stosich, E. (2016). *Pathways to new accountability through the Every Student Succeeds Act*. Learning Policy Institute.
- Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53(3), 285-328.
- Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher*, 31(9), 13-25.
- Davis, J. R., & Warner, N. (2015). Schools matter: The positive relationship between New York City high schools' student academic progress and school climate. *Urban Education*, 0042085915613544.
- Davis, L. (2016) Designing and using rubrics. In J. D. Brown & C. Coombe (Eds.), *The Cambridge guide to research in language teaching and learning* (238-246). Cambridge: Cambridge University Press.
- Doherty, K. M., & Jacobs, S. (2015). State of the states 2015: Evaluating teaching, leading and learning. *National Center on Teacher Quality*. Retrieved from the *National Center on Teacher Quality website*: <http://www.nctq.org/dmsView/StateofStates2015>
- Donaldson, M. (2009). So long Lake Wobegon? Using teacher evaluation to improve teacher quality. *Center for American Progress*.
- Downey, D. B., Von Hippel, P. T., & Hughes, M. (2008). Are "failing" schools really failing? Using seasonal comparison to evaluate school effectiveness. *Sociology of Education*, 81(3), 242-270.

- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101.
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91, 166-174.
- Duckworth, A. L., White, R. E., Matteucci, A. J., Shearer, A., & Gross, J. J. (2016). A stitch in time: Strategic self-control in high school and college students. *Journal of educational psychology*, 108(3), 329.
- Duckworth, A. L., Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher* 44(4), 237-251.
- Durlak, J. A., Domitrovich, C. E., Weissberg, R. P., & Gullotta, T. P. (Eds.). (2015). *Handbook of social and emotional learning: Research and practice*. New York, NY: Guilford.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrockk (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: MacMillan.
- Ericsson, A.K., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49, 725-747.
- Esch, C. E., Chang-Ross, C. M., Guha, R., Tiffany-Morales, J., & Shields, P.M. (2004). *California's teaching force 2004: Key issues and trends*. Santa Cruz, CA: The Center for the Future of Teaching and Learning.
- Evans, G. W. & Rosenbaum, J. (2008). Self-regulation and the income-achievement gap. *Early Childhood Research Quarterly*, 23, 504-514.

- Fadel, C., Bialik, M., & Trilling, B. (2015). *Four-dimensional education: The competencies learners need to succeed*. Boston, MA: The Center for Curriculum Redesign.
- Farnham, L., Fernando, G., Perigo, M., Brosman, C., & Tough, P. (2015). Rethinking how students succeed. *Stanford Social Innovation Review*. Retrieved from http://ssir.org/up_for_debate/article/rethinking_how_students_succeed
- Farrington, C.A., Roderick, M., Allensworth, E., Nagaoka, J., Keys, T.S., Johnson, D.W., & Beechum, N.O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance. A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.
- Ferguson, R. F, & Danielson, C. (2014). How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In T. J. Kane, K. A. Kerr & R. C. Pianta (Eds). *Designing teacher evaluation systems*. (pp. 98 - 143). San Francisco, CA: Jossey-Bass.
- Flynn, R. M., Lissy, R., Alicea, S., Tazartes, L., & McKay, M. M. (2016). Professional development for teachers plus coaching related to school-wide suspensions for a large urban school system. *Children and Youth Services Review*, 62, 29-39.
- Foster, C. (2016). Confidence and competence with mathematical procedures. *Educational Studies in Mathematics*, 91(2), 271-288.
- Fraser, B. J., & Walberg, H. J. (1991). *Educational environments: Evaluation antecedents and consequences*. Oxford: Pergamon Press.
- Gargani, J. & Strong, M. (2014). Can we identify a successful teacher better, faster, and cheaper? Evidence for innovating teacher observation systems. *Journal of Teacher Education*, 65(5), 389-401.

- Gist, M.E. & Mitchell, T.R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, 17(2), 183-211. Retrieved from <http://www.jstor.org.proxy.bsu.edu/stable/258770>
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Goddard, R. D. (2001). Collective efficacy: a neglected construct in the study of schools and student achievement. *Journal of Educational Psychology*, 93(3), 467-476.
- Goddard, R. D. (2002). Collective efficacy and school organization: a multilevel analysis of teacher influence in schools. *Theory and Research in Educational Administration*, 1, 169-184
- Goddard, R.D., & Goddard, Y.L. (2001). A multilevel analysis of the relationship between teacher and collective efficacy in urban schools. *Teaching and Teacher Education*, 17, 807-818.
- Goddard, R.D., Goddard, Y.L., Kim, E.S., & Miller, R. (2015). A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *American Journal of Education*, 121(4), 501-530.
- Goddard, R.D., Hoy, W.K., & Hoy, A.W. (2000). Collective teacher efficacy: Its meaning, measure and impact on student achievement. *American Educational Research Journal*, 37(2), 479-507.
- Goddard, Y. L., Goddard, R. D., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109, 877-896.

- Goetz, T., Nett, U. E., & Hall, N. C. (2013). Self-regulated learning. In Hall, N. C. & Gotez, T. (Eds.), *Emotion, motivation, and self-regulation: A handbook for teachers* (123-166). Bingley, UK: Emerald.
- Goldstein, J., & Noguera, P. A. (2006). A thoughtful approach to teacher evaluation. *Educational Leadership*, 63(6), 31-37.
- Goldring, E., Grissom, J. A., Rubin, M., Neumerski, C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value added principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher*, 44(2), 96-104.
- Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of school psychology*, 45(1), 3-19.
- Greene, J. C., & Caracelli, V. J. (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. San, Francisco: Jossey-Bass Publishers.
- Gullickson, A.R., & Howard, B.B. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2nd Ed.). Thousand Oaks, CA: Corwin Press.
- Hanson, E. M. (2003). *Educational administration and organizational behavior* (5th ed.). Boston: Allyn and Bacon.
- Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know*. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138.
- Harter, S. (1981). A model of mastery motivation in children: Individual differences and developmental change. In W. A. Collins (Ed.), *Aspects of the development of competences: Minnesota Symposia on child psychology*, 14 (pp. 215-255). Hillsdale, N.J.: Erlbaum.

- Heckman J. J., & Kautz, T. D. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451-464.
- Herold, B. (January 24, 2015) Is ‘grit’ racist? *Education Week*. Retrieved from http://blogs.edweek.org/edweek/DigitalEducation/2015/01/is_grit_racist.html.
- Hill, H. C., & Grossman, P. (2013). Learnings form teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review* 83(2), 371-384.
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hoy, W. K., Sweetland, S. R., & Smith, P. A. (2002). Toward an organizational model of achievement in high schools: The significance of collective efficacy. *Educational Administration Quarterly*, 38(1), 77-93.
- Hushman, C. J., & Marley, S. C. (2015). Guided instruction improves elementary student learning and self-efficacy in science. *The Journal of Educational Research*, 108(5), 371-381.
- Indiana Department of Education (IDOE). (2017). Indiana standards for mathematics – Fifth grade. Retrieved from <https://www.doe.in.gov/sites/default/files/standards/mathematics/grade-5-correlation-guide.pdf>.
- Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, 38(3), 499–534.
- Ingersoll, R. M., & Merrill, L. (2012). *Seven trends: The transformation of the teaching force*. The Consortium for Policy Research in Education.

- John, O. P. & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York, NY: Guilford Press.
- Johnson, S. M., Berg, J. H., & Donaldson, M. L. (2005). *Who stays in teaching and why? A review of the literature on teacher retention*. Washington, DC: National Retired Teachers Association.
- Judge, S., & Watson, S. M. R. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. *Journal of Educational Research*, 104, 147–157.
- Kalogrides, D., Loeb, S., & Bételle, T. (2012). Systematic sorting: Teacher characteristics and class assignments. *Sociology of Education* 86(2), 103-1123.
- Kane, T. J., & Staiger, D. (2008). Estimating teacher impacts on student achievement: An experimental evaluation (NBER Working Paper 14607). Cambridge, MA: National Bureau of Economic Research.
- Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have we identified effective teachers? Validating measures of effective teaching using random assignment. In *Research Paper. MET Project. Bill & Melinda Gates Foundation*.
- Karwowski, M. (2013). Creative mindsets: Measurement, correlates, and consequences. *Psychology of Aesthetics, Creativity and the Arts*, 8(1), 62-70.
- Kast, F., & Rosenzweig, J. (1972). General systems theory: Applications for organization and management. *The Academy of Management Journal*, 15(4), 447-465.
- Krachman, S. B., Arnold, R., Larocca, R. (September 2016). Expanding the definition of student success: A case study of the CORE districts. Retrieved from <http://www.transformingeducation.org/core-toolkit/>.

- Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment?. *Child development, 61*(4), 1081-1100.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.
- Lin-Siegler, X., Dweck, C. S., & Cohen, G. L. (2016). Instructional interventions that motivate classroom learning. *Journal of Educational Psychology, 108*(3), 295.
- Lincoln, R. D. (1997). Multi-year instruction: Establishing student-teacher relationships. *Schools in the Middle, 6*(3), 50-52.
- Lodewyk, K. R., & Winne, P. H. (2005). Relations among the structure of learning tasks, achievement, and changes in self-efficacy in secondary students. *American Psychological Association, 97*(1), 2-12.
- Lüftenegger, M., Finsterwald, M., Klug, J., Bergsmann, E., van de Schoot, R., Schober, B., & Wagner, P. (2016). Fostering pupils' lifelong learning competencies in the classroom: evaluation of a training programme using a multivariate multilevel growth curve approach. *European Journal of Developmental Psychology, 13*(6), 719-736.
- Lyman, R. D. (1984). The effect of success or failure on self-efficacy and task persistence of conduct-disordered children. *Psychology in the Schools, 21*, 516-519.
- Manstead, A. S. R., van Eekelen, S. A. M (1998). Distinguishing between perceived behavioral control and self-efficacy in the domain of academic achievement intentions and behaviors. *Journal of Applied Social Psychology, 28*(15), 1375-1392.

- Marinell, W. H., & Coca, V. M. (2013). *Who stays and who leaves? Findings from a three part study of teacher turnover in NYC middle schools*. New York: The Research Alliance for NYC Schools.
- Marshik, T., Ashton, P. T., & Algina, J. (2017). Teachers' and students' needs for autonomy, competence, and relatedness as predictors of students' achievement. *Social Psychology of Education, 20*(1), 39-67.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19.
- Marzano, R. J. Carbaugh, B., Rutherford, A., Toth, M. D. (2014). *Marzano center observation protocol for the 2014 Marzano teacher evaluation model*. Retrieved from http://www.marzanocenter.com/files/2014_Protocol_Paper_20140128.pdf.
- Marzano, R. J. & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth*. Alexandria, VA: ASCD.
- Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American psychologist, 53*(2), 205.
- Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior, 37*(2), 225-238.
- Measures of Effective Teaching; Bill and Melinda Gates Foundation. Retrieved from http://www.hec.ca/iea/seminaires/140401_staiger_douglas.pdf.
- Mega, C., Ronconi, L., De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology 106*(1), 121-131.

- Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.
- Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. [Research manual]. University of Michigan.
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis*. Thousand Oaks: Sage Publications.
- Mirowsky, J., & Ross, C. E. (2003). *Social causes of psychological distress*. New York, NY: Walter de Gruyter
- Moolenaar, N. M., Slegers, P. J. C., & Daly, A. J. (2011). Teaming up: Linking collaboration networks, collective efficacy, and student achievement. *Teaching and Teacher Education, 28*, 251-262.
- Mone, M A., Baker, D. D., & Jeffries, F. (1995). Predictive validity and time dependency of self-efficacy, self-esteem, personal goals, and academic performance. *Educational and Psychological Measurement, 55*(5), 716-727.
- Moore, T. T., Chang, J. C., Smith, D. K. (2006). Clarifying the role of self-efficacy and metacognition as indicators of learning: Construct development and test. *The DATA BASE for Advances in Information Systems, 37*(2 & 3), 125-132. Retrieved from http://digitalscholarship.unlv.edu/cgi/viewcontent.cgi?article=1007&context=met_fac_articles
- Moos, D. C. (2014). Setting the stage for the metacognition during hypermedia learning: What motivation constructs matter? *Computers & Education, 70*, 128-137.

- Morse, J. M. (1991). Strategies for sampling. In J. M. Morse (Ed.), *Qualitative nursing research: A contemporary dialogue* (pp. 127-145). Thousand Oaks, CA, US: Sage Publications, Inc.
- Morse, J.M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189-208). Thousand Oaks, CA: Sage.
- Mouratidis, A., Vansteenkiste, M., Michou, A., Lens, W. (2013). Perceived structure and achievement goals as predictors of students' self-regulated learning and affect and the mediating role of competence need satisfaction. *Learning and Individual Differences*, 23, 179-186.
- Moustakas, Clark E. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage Publications, Inc.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30-38.
- Mullainathan, S. & Shafir, E. (2013). *Scarcity: Why having too little means so much*. New York, NY: Time Books.
- National Institute for Excellence in Teaching (1999). *TAP rubric*.
- Onwuegbuzie, A. J., & Collins, K. M. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, 12(2), 281-316.
- Onwuegbuzie, A. J., Jiao, Q. G., & Bostick, S. L. (2004). Library anxiety: Theory, research, and applications. Lanham, MD: Scarecrow Press.

- Ovando, M. N. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *Journal of Personnel Evaluation in Education, 15*(3), 213-231.
- Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research, 66*(4), 543-578.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Reading & Writing Quarterly, 12*, 139-158.
- Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research and applications*. New York: Erlbaum.
- Pajares, F. & Johnson, M. (1994). Confidence and competence in writing: the role of self-efficacy, outcome expectancy and apprehension. *Research in the Teaching of English, 28*(3), 313-331.
- Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English, 104*-120.
- Pajares, F. & Miller, M. D. (1995). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*(2), 193-203.
- Papay, J. (2012) Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review, 82*(1), 123-141.
- Paunesku, D. Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 1*-10.

- Phan, H. P. (2013). Examination of self-efficacy and hope: A developmental approach using latent growth modeling. *The Journal of Educational Research* 106(2), 93-104.
- Pianta, R. C., La Paro, K., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Paul H. Brookes.
- Pina-Neves S., Faria, L., & Ráty, H. (2013). Students' individual and collective efficacy: joining together two sets of beliefs for understanding academic achievement. *European Journal of Psychology of Education*, 28(2), 453-474.
- Pintrich, P. R., & De Groot, E. V., (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.
- Putney, L. G. (1996). You are it: Meaning making as a collective and historical process. *Australian Journal of Language and Literacy*, 19, 129-143.
- Putney, L.G. & Broughton, S.H. (2011). Developing collective classroom efficacy: The teacher's role as community organizer. *Journal of Teacher Education*, 62(1), 93-105.
- Quiroga, C. V., Janosz, M., Bisset, S., & Morin, A. J. S. (2013, February). Early adolescent depression symptoms and school dropout: Mediating processes involving self-reported academic competence and achievement. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/a0031524.
- Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science*, 10(6), 721-726.
- Reardon, S.F., (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. Murnane & G. Duncan (Eds), *Whither*

- Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children* (pp. 91-115). New York: Russell Sage Foundation Press.
- Reeve, J., Ryan, R., Deci, E. L., & Jang, H. (2008). Understanding and promoting autonomous self-regulation: A self-determination theory perspective. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 223-244). New York, NY: Lawrence Erlbaum.
- Reform Support Network. (July, 2015). *Using observations to improve teacher practice*. Retrieved from <https://www2.ed.gov/about/inits/ed/implementation-support-unit/tech-assist/usingobservations toimproveteacherpractice.pdf>
- Ris, E. (2015). Grit: A short history of a useful concept. *Journal of Educational Controversy*, 10(1). Retrieved from <http://cedar.wvu.edu/jec/vol10/iss1/3>.
- Ritter, G. W. & Barnett, J. H. (2016). Learning on the job: Teacher evaluation can foster real growth. *Phi Delta Kappan*, 97(7), 48-52.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrics*, 73(2), 417-458.
- Roderick, M., Coca, V., & Nagaoka, J. (2011). Potholes on the road to college: High school effects in shaping urban students' participation in college application, four-year college enrollment, and college match. *Sociology of Education*, 84, 178-211
- Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4-36.
- Rosenholtz, S. J., & Simpson, C. (1984). The formation of ability conceptions: Developmental trend or social construction?. *Review of Educational Research*, 54(1), 31-63.

- Rosenholtz, S. J., & Wilson, B. (1980). The effect of classroom structure on shared perceptions of ability. *American Educational Research Journal*, 17(1), 75-82.
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics*, 125(1), 175–214.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(1), 1-28.
- Rutledge, S. A, Cohen-Vogel, L., Osborne-Lampkin, L., Roberts, R. L. (2015). Understanding effective high schools: Evidence for personalization for academic and social emotional learning. *American Educational Research Journal*, 52(6), 1060-1092.
- Ryan, A. M., Gheen, M. H., & Midgley C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social-emotional role, and the classroom goal structure. *Journal of Educational Psychology*, 90(3), 528-535.
- Saldaña, J. (2013). *The coding manual for qualitative researchers*. Los Angeles: SAGE Publications.
- Sartain, L., Stoelinga, S.R., & Brown, E.R. (2011). Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation. *Consortium on Chicago School Research*.
- Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*. 73: 93-105.
- Schunk, D. H. (1984). Sequential attributional feedback and children's achievement behaviors. *Journal of Educational Psychology*, 76, 1159-1169.

- Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of educational research*, 57(2), 149-174.
- Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. In C. Ames and R. Ames (Eds.), *Research on motivation in education. Goals and cognitions*, 3 (pp. 13-44). San Diego: Academic.
- Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational psychologist*, 25(1), 71-86.
- Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application* (pp. 281-303). New York: Plenum Press.
- Schunk, D. H., & Hanson, A. R. (1989). Influence of peer-model attributes on children's beliefs and learning. *Journal of Educational Psychology*, 81, 431-434.
- Schunk, D. H., & Rice, J. M. (1987). Enhancing comprehension skill and self-efficacy with strategy value information. *Journal of Reading Behavior*, 19(3), 285-302.
- Schunk, D.H., & Zimmerman, B.J. (Eds.). (1998). *Self regulated learning: From teaching to self-reflective practice*. New York: Guilford Press.
- Schwartz, D., Gorman, A. H., Nakamoto, J., & Toblin, R. L. (2005). Victimization in the peer group and children's academic functioning. *Journal of educational psychology*, 97(3), 425.
- Sheridan, K. M., Clark, K., & Williams, A. (2013). Designing games, designing roles: A study of youth agency in urban informal education program. *Urban Education* 48(5). 734-758.
- Schulte, A. C. & Stevens, J. J. (2015). Once, sometimes, or always in special education: Mathematics growth and achievement gaps. *Exceptional Children*, 81(3), 370-387.

- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), p. 153-189.
- Socol, I. (2014). Taking a closer look at the grit narratives. *Knowledge Quest* 43(1), p. 8-12.
- Soland, J. Hamilton, L. S., & Stecher, B. M. (2013). *Measuring 21st century competencies: Guidance for educators*. Santa Monica, CA: RAND Corporation.
- Spradlin, T. E., Dickinson, S. L., Dingjing, L., Shi, D., & Whiteman, R. S. (2012). *Acuity Assessment Program Indiana Efficacy Study: 2008-2011 Data Analysis (Final Edition)*. Bloomington, IN: Center for Evaluation & Education Policy.
- St Clair-Thompson, H., Bugler, M., Robinson, J., Clough, P., McGeown, S. P., & Perry, J. (2015). Mental toughness in education: Exploring relationships with attainment, attendance, behaviour and peer relationships. *Educational Psychology*, 35(7), 886-907.
- Stecher, B. M., & Hamilton, L. S. (2014). *Measuring hard-to-measure student competencies: A research and development plan*. Santa Monica, CA: RAND Corporation.
- Steinberg, M., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis*. 38(2), 293-317.
- Steiner, L. (2010). Using competency based evaluation to drive teacher excellence: Lessons from Singapore. *Public Impact*. Retrieved from <http://files.eric.ed.gov.proxy.bsu.edu/fulltext/ED539994.pdf>.
- Strauss, A., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage Publications.
- Strong, M. (2011). *The highly qualified teacher. What is teacher quality and how do we measure it?* New York: Teachers College Press.

- Stuhlman, M. W., & Pianta, R. C. (2009). Profiles of educational quality in first grade. *The Elementary School Journal*, 109(4), 323-342.
- Tashakkori, A., Teddlie, C., & Teddlie, C.B. (Eds.). (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, Calif: SAGE Publications.
- Taylor, Eric S., & Tyler, J. H. (2012a). The effect of evaluation on teacher performance. *American Economic Review*, 102(7), 3628-3651.
- Taylor, E. S., & Tyler, J. H. (2012b) Can teacher evaluation improve teaching? *Education Next* 12(4), 78-84.
- Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research*, 27, 293–302.
- Tough, P. (2015). Response to “rethinking how students succeed”. *Stanford Social Innovation Review*. Retrieved from http://ssir.org/up_for_debate/rethinking_how_students_succeed/paul_tough
- Tough, P. (2013). *How children succeed: Grit, curiosity, and the hidden power of character*. New York, NY: Houghton Mifflin Harcourt.
- Transforming Education (April, 2016). Measuring MESH: Student and teacher surveys curated for the CORE districts. Retrieved from <http://www.transformingeducation.org/measuringmesh/>.
- Tschannen-Moran, M., Woolfold Hoy, A., & Hoy, W. K. (1998). Teacher Efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202-248.
- Tsukayama, E., Duckworth, A.L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, 16, 879-893.

- Tulis, M., Und, M. R., & Dresel, M. (2018). Friendships in the classroom and their importance for adaptive individual dealing with errors. *Zeitschrift Fur Entwicklungspsychologie Und Padagogische Psychologie*, 50(1), 44-58.
- US Department of Education. (2004). New no child left behind flexibility: Highly qualified teachers. Retrieved from <http://www2.ed.gov/nclb/methods/teachers/hqtflexibility.pdf>.
- Usher, E. L (2009). Sources of middle school students' self-efficacy in mathematics: A qualitative investigation. *American Educational Research Journal*, 46(1), 275-314.
- Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology*, 31(2), 125-141.
- Vancouver, J.B., Thompson, C.M., Tischner, E.C., & Putka, D.J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology*, 87(3), 506-516.
- Vanderburg, M. & Stephens, D. (2010). The impact of literacy coaches: What teachers value and how teachers change. *The Elementary School Journal*, 111(1), 141-163.
- Webb, N. L. (1997). Criteria for alignment of expectations and assessments on mathematics and science education. *National Institute for Science Education*. Washington, D.C.: CCSSO.
- Webb, N.L. (2002). *Depth-of-knowledge levels for the four content areas*. Retrieved from <http://ossucurr.pbworks.com/w/file/fetch/49691156/Norm%20web%20dok%20by%20subject%20area.pdf>.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. *The Education Digest*, 75(2), 31-35.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child development*, 62(5), 1066-1078.

- West, M. R. (2014). *Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling*. Retrieved from <http://cepr.harvard.edu/files/cepr/files/cepr-promise-paradox.pdf>.
- West, M. R. (2016). *Should non-cognitive skills be included in school accountability systems? Preliminary evidence from California's CORE districts*. Evidence Speaks Reports, 1(13).
- West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, 38(1), 148-170.
- White, R. W. (1959). Motivation reconsidered: The concept of competences. *Psychological Review*, 66, 297-333.
- White, R. W. (1960). Competence and the psychosexual stages of development. In M.R. Jones (Ed.), *Nebraska Symposium on Motivation*, 8, (pp. 97-141). Lincoln: University of Nebraska Press.
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91, 301-311.
- Wright, S. P., Horn, S., P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education* 11, 57-67.
- Wood, R. & Bandura, A. (1989). Social cognitive theory of organizational management. *The Academy of Management Review*, 14(3), 361-384.

- Yerdelen, S., & Sungur, S. (2019). Multilevel investigation of students' self-regulation processes in learning science: Classroom learning environment and teacher effectiveness. *International Journal of Science and Mathematics Education*, 17(1), 89-110.
- Yarrow, L. J., McQuiston, S., MacTurk, R. H., McCarthy, M. E., Klein, R. P., & Vietze, P. M. (1983). Assessment of mastery motivation during the first year of life: Contemporaneous and cross-age relationships. *Developmental Psychology*, 19(2), 159-171.
- Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). Practical measurement. *Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching*.
- Zahodne, L. B., Nowinski, C. J., Gershon, R. C., & Manly, J. J. (2015). Self-efficacy buffers the relationship between educational disadvantage and executive functioning. *Journal of the International Neuropsychological Society*, 21(04), 297-304.
- Zimmerman, B. J. (1999). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82-91.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner. *Theory Into Practice* 41(2), 64-70.
- Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performances. In B. J. Zimmerman & D. H. Schunk (Eds.) *Handbook of self-regulation of learning and performance* (pp. 49-64). New York: Routledge.
- Zimmerman, B.J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of Educational Psychology*, 91, 1-10

APPENDIX A

The following survey was developed and validated by the Core Districts (2016) to collect MESH data on students. Question 10 was modified from “I can earn an A in my classes” to “I can good grades in my classes” to account for the untraditional grading practices of the schools.

Self-Management:

First, we’d like to learn more about your behavior, experiences, and attitudes related to school.

Please answer how often you did the following during the past 30 days. During the past 30 days...

(Almost Never, Once in a While, Sometimes, Often, Almost All the Time)

1. I came to class prepared.
2. I remembered and followed directions.
3. I got my work done right away instead of waiting until the last minute.
4. I paid attention, even when there were distractions.
5. I worked independently with focus.
6. I stayed calm even when others bothered or criticized me.
7. I allowed others to speak without interruption.
8. I was polite to adults and peers.
9. I kept my temper in check.

Self-Efficacy:

How confident are you about the following at school?

(Not At All Confident, A Little Confident, Somewhat Confident, Mostly Confident, Completely Confident)

10. I can get good grades in my classes.
11. I can do well on all my tests, even when they’re difficult.
12. I can master the hardest topics in my classes.
13. I can meet all the learning goals [targets] my teachers set.

Social Awareness:

In this section, please help us better understand your thoughts and actions when you are with other people.

Please answer how often you did the following during the past 30 days. During the past 30 days...

14. How carefully did you listen to other people's points of view?

(Not Carefully At All, Slightly Carefully, Somewhat Carefully, Quite Carefully, Extremely Carefully)

15. How much did you care about other people's feelings?

(Did Not Care At All, Cared A Little Bit, Cared Somewhat, Cared Quite A Bit, Cared A Tremendous Amount)

16. How often did you compliment others' accomplishments?

(Almost Never, Once in a while, Sometimes, Often, Almost all the time)

17. How well did you get along with students who are different from you?

(Did Not Get Along At All, Got Along A Little Bit, Got Along Somewhat, Got Along Pretty Well, Got Along Extremely Well)

18. How clearly were you able to describe your feelings?

(Not At All Clearly, Slightly Clearly, Somewhat Clearly, Quite Clearly, Extremely Clearly)

29. When others disagreed with you, how respectful were you of their views?

(Not At All Respectful, Slightly Respectful, Somewhat Respectful, Quite Respectful, Extremely Respectful)

20. To what extent were you able to stand up for yourself without putting others down?

(Not At All, A Little Bit, Somewhat, Quite A Bit, A Tremendous Amount)

21. To what extent were you able to disagree with others without starting an argument?

(Not At All, A Little Bit, Somewhat, Quite A Bit, A Tremendous Amount)

APPENDIX B

The following questions extracted from the PALS survey, comprising the subscales of Teacher Mastery Goal and Academic Press.

This permission is granted on the condition that LICENSEE properly attributes the MANUAL to the University of Michigan and includes the following copyright notice "Copyright © 2000 The Regents of the University of Michigan."

Academic Press:

1. When I've figured out how to do a problem, my teacher gives me more challenging problems to think about.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

2. My teacher presses me to do thoughtful work.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

3. My teacher asks me to explain how I get my answers.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

4. When I'm working out a problem, my teacher tells me to keep thinking until I really understand.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

5. My teacher doesn't let me do just easy work, but makes me think.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

6. My teacher makes sure that the work I do really makes me think.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

7. My teacher accepts nothing less than my full effort.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

Teacher Mastery Goal:

8. My teacher thinks mistakes are okay as long as we are learning.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

9. My teacher wants us to understand our work, not just memorize it.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

10. My teacher really wants us to enjoy learning new things.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

11. My teacher recognizes us for trying hard.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

12. My teacher gives us time to really explore and understand new ideas.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

APPENDIX C

The following questions were extracted from the CORE District survey (2016) and were administered to teachers upon the conclusion of the study.

1. Please consider the full set of behaviors and provide a single overall rating for this student. In other words, you will provide three ratings for the student: one for self-management (school work), one for self-management (interpersonal), and one for social awareness. You do NOT need to rate this student on each individual behavior listed.

Self-Management (School Work) - How often did this student exhibit the following behaviors:

- Came to class ready to learn.
- Remembered and followed directions.
- Persisted when tasks became challenging.
- Paid attention and maintained focus.
- Resisted distractions.

1	2	3	4	5
Almost never	Once in a while	Sometimes	Often	Almost all of the time

2. Self-Management (Interpersonal) - How often did this student exhibit the following behaviors:

- Remained calm even when under stress.
- Allowed others to speak without interruption.
- Got along well with others.
- Kept his/her temper in check.

1	2	3	4	5
Almost never	Once in a while	Sometimes	Often	Almost all of the time

3. Social Awareness - How often did this student exhibit the following behaviors:

- Listened carefully to other people's points of view.
- Got along with students who were different from him/her.

- Disagreed with someone without starting an argument.
- Stood up for him/herself without putting others down.
- Noticed and complimented others' accomplishments.

1	2	3	4	5
Almost never	Once in a while	Sometimes	Often	Almost all of the time

APPENDIX D

Tool 1: The following is the first draft of the observation data collection form constructed after one round of observation and open coding.

Teacher:	Subject:	Date:	Time:
Student Behaviors	# of Instances	Notes	
Evidence that students manage their time effectively *and have an appropriate amount of time to explore ideas of the content.			
Evidence that students transfer knowledge and skills and select appropriate strategies to apply these in new situations.			
Evidence that students demonstrate perseverance and problem solving, including asking for help, when faced with barriers to success.			
Evidence that students don't reject good solutions prematurely.			
Evidence that students take academic risks by taking on more challenging tasks, setting higher standards and setting more challenging goals for themselves (choice in how they complete tasks as well).			
Evidence that students track, achieve and *celebrate reaching their goals.			
Evidence that students use more efficient strategies.			
Evidence that students incorporate self-reflection, metacognition, and growth-focused language while owning their decisions and the consequences via strategic thinking.			
Evidence that students use metacognition to monitor their understanding and the relevance of the content.			
Evidence that students self-regulate during the process of learning by monitoring their own effort, work			

quality, and learning habits.		
*Evidence that students are able to confidently agree or disagree with each other about content		
*Evidence that students are expected to think about / engage with rigorous, grade-appropriate content		
*Evidence that students are able to use the tools of the content, including vocabulary.		
Intentionality – action plan for goals		
Forethought – visualize future & outcomes		
Self-reaction – self-regulation		
Self-reflection		
Mastery Experiences		
Vicarious experiences		
Verbal persuasion		
Physiological & affective states		

APPENDIX E

Tool 2: The following is the second iteration of the observation tool reflecting the Eclectic and Axial Coding process of round one and two data.

<p>Experience success</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Use peer models for guidance</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Critique work and give feedback toward the goal</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>
<p>Think about / engage with rigorous, grade-appropriate content (do the cognitive work)</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Use peer models for self-evaluation or comparison</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Assess academic work</p> <p>Students independently: NA None Little Some Most All</p> <p>Students with teacher support or prompting: NA None Little Some Most All</p> <p>Teacher: NA None Little Some Most All</p>
<p>Utilize prior knowledge or strategies to persevere through work when challenges arise</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Respectfully share thinking and differing opinions</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting: None Few Some Most All</p> <p>Teacher: None Few Some Most All</p>	<p>Assess work behaviors/ effort</p> <p>Students independently: NA None Little Some Most All</p> <p>Students with teacher support or prompting: NA None Few Some Most All</p> <p>Teacher: NA None Few Some Most All</p>
<p>Appraise and monitor progress</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting:</p>	<p>Interact with peers to support or complete work</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting:</p>	<p>Celebrate or offer encouragement for academic accomplishments and/or hard work</p> <p>Students independently: None Few Some Most All</p> <p>Students with teacher support or prompting:</p>

None Few Some Most All	None Few Some Most All	None Few Some Most All
Teacher: None Few Some Most All	Teacher: None Few Some Most All	Teacher: None Few Some Most All
Regulate emotion when it comes to unsuccessful experiences	Regulate / manage behaviors	Manage time effectively / are productive
Students independently: None Few Some Most All	Students independently: None Few Some Most All	Students independently: None Few Some Most All
Students with teacher support or prompting: None Few Some Most All	Students with teacher support or prompting: None Few Some Most All	Students with teacher support or prompting: None Few Some Most All
Teacher: None Few Some Most All	Teacher: None Few Some Most All	Teacher: None Few Some Most All

APPENDIX F

Tool 3: The following is the final iteration of the observation tool developed and used for this study.

	Task Experience / Responsibility Scale: 3 - Students independently of teacher 2 - Students with teacher support or prompting 1 - Mostly or all teacher NO – Not Observed					Notes
Mastery Experiences	Experience Success	3	2	1	NO	
	Think about / engage with rigorous, grade-appropriate content (do the cognitive work)	3	2	1	NO	
	Utilize prior knowledge or strategies to persevere through work when challenges arise	3	2	1	NO	
	Appraise and monitor progress	3	2	1	NO	
Vicarious Experiences	Use peer models for guidance	3	2	1	NO	
	Use peer models for self-evaluation or comparison	3	2	1	NO	
	Respectfully share thinking and differing opinions	3	2	1	NO	
	Interact with peers to support or complete work	3	2	1	NO	
Verbal Persuasion	Critique work and give feedback toward the goal	3	2	1	NO	
	Assess academic work	3	2	1	NO	
	Assess work behaviors	3	2	1	NO	
	Celebrate or offer encouragement for academic and/or CHOW accomplishments	3	2	1	NO	
Physiological & Affective States	Regulate emotion when it comes to unsuccessful experiences	3	2	1	NO	
	Regulate / manage behaviors	3	2	1	NO	
	Manage time effectively / are productive	3	2	1	NO	

APPENDIX G

Sample scoring look-fors used in training and inter-rater reliability checks.

Source category subscale	Indicator	Sample Scoring Look-Fors
There is evidence that...		
3 - Students independently of teacher 2 - Students with teacher support or prompting 1 - Mostly or all teacher NO – Not Observed		
Mastery Experiences	Experience Success	(3 or 2) Students are able to complete the task and/or learning/CHOW target successfully; acquire knowledge or develop the skills of the content needed to accomplish the task/target successfully, indicate they feel they can do the work successfully. (2) Teacher may model the task and students emulate the teacher. Students demonstrate or indicate they could do the task/target independently. (1) The teacher asks students to copy work and/or students indicate they believe they cannot do the work independent of the teacher after a model. (NO) The majority of students are not successfully able to complete the task, or indicate that they feel unsuccessful or confused.
	Think about / engage with rigorous, grade-appropriate content (do the cognitive work)	(3 or 2) Students are asked to actively complete work and/or think about/ grapple with content aligned to a grade-level state standard. In math this includes the aspect of rigor, and in ELA the text is appropriately complexity for the grade level. The work is also aligned to the appropriate DOK level tending to Levels 2-4. (1) The teacher models how to do the grade-level aligned work (with alignment to rigor, text complexity, and/or DOK) but the students are mostly passive or watch the teacher work. (NO) Students complete work unaligned grade-level standards, the DOK remains inappropriately at the Level 1, the aspect of rigor in math is misaligned to the standard, and/or the text complexity is below grade-level in ELA. This could also be used to indicate students are working on content skills outside of content (such as CHOW).
	Utilize prior knowledge or strategies to persevere through work when challenges arise	(3 or 2) Students use strategies such as sounding out unknown words, using the web or other tools, relying on conceptual understanding to figure out the procedure, etc. For CHOW, this could look like using conflict resolution skills, organizing skills, etc. to problem solve. (1) The teacher prompts students to use specific strategies to overcome their challenge. The teacher may rescue students when they are stuck, prompt them through the process, or show them how to do the task, etc. (NO) Students do not encounter challenges or give up when they encounter them with no specific guidance or intervention on techniques/strategies from the teacher.

	Appraise and monitor progress	<p>(3) Students self-assess their progress towards success on the task, targets, or behavior objectives, against quality indicators, exemplars, etc. without or with minor support from the teacher.</p> <p>(1) The teacher tells the student how they are progressing on the task, targets, or behavior objectives or students are simply told to get to work, complete the task differently, etc.</p> <p>(NO) Students do not self-monitor or appraise their progress towards the task nor do they receive comparable or effective feedback from the teacher.</p>
Vicarious Experiences	Use peer models for guidance	<p>(3 or 2) Students look to other students or their work for guidance on what to do, how to solve a problem, how to revise their thinking, how to behave, etc. (i.e. develop strategy knowledge)</p> <p>(2) The teacher may point out behaviors in students for others to emulate, or use a student or their work in class as a model or exemplar for the task, target, or behavior, etc.</p> <p>(1) The teacher shows students what to do for the task or how to behave, etc.</p> <p>(NO) The task and/or behavior were not modeled.</p>
	Use peer models for self-evaluation or comparison	<p>(3 or 2) Students use each other or a comparison of their work to determine their level of mastery, achievement, success, etc.</p> <p>(1) The teacher tells the students how they are doing relative to one another, etc. or the self-evaluation is so generic that it serves as a simple data point, rather than inform their self-evaluation.</p> <p>(NO) The students are not aware or do not engage in activities that give them an idea of how they are doing relative to their peers, etc.</p>
	Respectfully share thinking and differing opinions	<p>(3 or 2) Students collaborate and share their thinking with each other respectfully, disagree without anger or conflict, etc.</p> <p>(1) The teacher calls on students and moderates/controls the conversation, and/or controls the way students hear or interpret other students' opinions</p> <p>(NO) The students aren't respectful and the teacher doesn't control or temper conversations, or students aren't given the opportunity to share them.</p>
	Interact with peers to support or complete work	<p>(3) Students are free to use peers when and how they want to complete the task or learn content and take the opportunity to do so.</p> <p>(2) The teacher prompts or dictates interactions between students and/or the topics when completing the task or learning content.</p> <p>(1) The teacher is only the to interact with and support students during the work completion and learning process.</p> <p>(NO) Most students work independently and do not interact with anyone to complete their work or learn the content.</p>
Verbal Persuasion	Critique work and give feedback toward the goal	<p>(3 or 2) Students reflect on the quality of work as it relates to the LT, give feedback, or use steps in feedback protocols on their goals or progress towards meeting their learning targets, CHOW targets, etc.</p> <p>(1) The teacher is the one who reflects and gives students feedback and/or critique towards the learning or CHOW target or goal.</p>

		(NO) There is no evidence that students or teachers engage in giving feedback towards goals learning targets, or CHOW targets and/or grades are given holistically for tasks/assignments.
	Assess academic work	(3 or 2) Students look for right and wrong “answers” (or application of skills in writing) within the task or work. They make value or quality judgments about their own work and set goals for their academic achievement. (1) The teacher tells students they have correct or incorrect “answers” (or application of skills in writing). The teacher gives students their goals for academic achievement. (NO) There is no evidence that students know if they have correct or incorrect “answers” (or application of skills in writing), or about the quality of their work. There don’t appear to be goals established for academic achievement.
	Assess work behaviors	(3 or 2) Students rate, score, or otherwise assess themselves on work behaviors/CHOW. Students set behavior/CHOW goals for themselves. (1) The teacher tells students how they are presenting work behaviors/CHOW. Teachers set behavior/CHOW goals for the students. (NO) There is no evidence that students know effective work behaviors and/or there are no behavior/CHOW goals established.
	Celebrate or offer encouragement for academic and/or CHOW accomplishments	(3 or 2) Students engage in celebratory behaviors for themselves or each others’ academic and/or CHOW accomplishments (such as cheering, saying good job, etc.) (1) The teacher is responsible for encouraging students and celebrating their successes and effective behaviors. Students are not engaged in this process. (NO) There is no evidence that teachers or students celebrate accomplishments or effective behaviors, or effort is acknowledged without connection to the specific behaviors that produce success.
	Physiological & Affective States	
	Regulate emotion when it comes to unsuccessful experiences	(3) When students get answers or processes wrong, they either don’t get frustrated or get frustrated but then calm themselves down to continue persevering through the task. (2) When students get answers or process wrong, they receive some minor teacher encouragement and calm down to continue persevering through the task (1) When students get answers or processes wrong, they receive major teacher support, such as redirection, teacher intervention to remove the barrier / frustration, or take the responsibility for calming down away from the student (such as a timeout from class or changing the task completely). (NO) There is no evidence that the task created frustration, confusion, etc. or that the students were able to handle it if present (such as students collectively getting angry, refusing to work, etc.)
	Regulate / manage behaviors	(3) Students regulate academic behaviors toward the task without needing reminders from the teacher but may receive reminders from peers (such as working at an appropriate pace, staying on task, working

	<p>alone or with peers, etc.)</p> <p>(2) Students need reminders the teachers for regulating their academic behaviors (such as working at an appropriate pace, staying on task, working alone or with peers, etc.)</p> <p>(1) Students are unable to regulate behaviors so the teacher makes choices that regulate them for them (such as changing group work to individual work to get students to stay on task) or removing students from the group setting (such as isolating students from group work, removing a student from the classroom, etc.)</p> <p>(NO) Students are completely off task, not following the stated or task-appropriate behavior expectations, etc. and the teacher doesn't intervene.</p>
Manage time effectively / are productive	<p>(3 or 2) Students work at an appropriate pace for the task. Time isn't wasted, nor do they work so fast that they make silly errors or take shortcuts in their work.</p> <p>(1) The teacher highly regulates the time using timers, etc. and moves the class on from task to task regardless of student need.</p> <p>(NO) The students do not manage their time effectively and there is no evidence that the teacher intervenes, or s/he uses management techniques such as timers, progress checks, etc. ineffectively so that time wasted or rushed to the point of the class being unproductive.</p>

APPENDIX H

The following tables present the significant and non-significant demographic variables relative to student outcomes and survey responses.

Table 18
Means, Standard Deviations, and Confidence Intervals of Acuity A outcomes in ELA and Mathematics by Demographic Variable

Group ^a	ELA			Mathematics		
	<i>M (SD)</i>	95% CI		<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>		<i>LL</i>	<i>UL</i>
Male ^b	.01 (.98)	-.28	.29	.11 (1.04)	-.19	.41
Female ^c	-.01 (1.01)	-.25	.24	-.08 (.96)	-.31	.16
Looped ^d	.26 (.98)*	-.04	.56	.36 (1.01)**	.05	.68
Did Not Loop ^e	-.15 (.97)*	-.38	.08	-.21 (.93)**	-.43	.01
White ^f	.03 (1.10)	-.48	.55	.02 (1.07)	-.48	.52
Non-White ^g	-.01 (.97)	-.21	.19	-.01 (.98)	-.21	.20
Paid Lunch ^h	-.14 (1.04)	-1.79	1.52	-.41 (1.43)	-2.68	1.86
FRL Lunch ⁱ	.01 (.99)	-.18	.19	.01 (.98)	-.17	.20
ELL ^j	.06 (.85)	-.22	.35	.01 (.88)	-.29	.30
Non-ELL ^k	-.03 (1.05)	-.27	.21	.00 (1.04)	-.24	.23
SPED ^l	-.57 (.64)*	-.93	-.20	-.44 (.91)	-.96	.08
Non-SPED ^m	.08 (1.01)*	-.12	.28	.06 (.99)	-.14	.26
Garfield ⁿ	.23 (1.01)**	-.03	.49	.33 (1.06)***	.06	.60
Washington ^o	-.28 (.90)**	-.53	-.03	-.40 (.74)***	-.60	-.19

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

Table 19

Means, Standard Deviations, and Confidence Intervals of Acuity B Outcomes in ELA and Mathematics by Demographic Variable

Group ^a	ELA			Mathematics		
	<i>M (SD)</i>	95% CI		<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>		<i>LL</i>	<i>UL</i>
Male ^b	.09 (.97)	-.19	.37	.22 (.73)*	.01	.43
Female ^c	-.07 (.96)	-3.1	.16	-.16 (1.12)*	-.44	.11
Looped ^d	.32 (1.02)**	.00	.64	.30 (.89)*	.02	.58
Did Not Loop ^e	-.19 (.93)**	-.41	.03	-.17 (1.01)*	-.41	.06
White ^f	.08 (1.18)	-.48	.63	.13 (.88)	-.28	.54
Non-White ^g	-.02 (.95)	-.21	.18	-.03 (1.102)	-.24	.18
Paid Lunch ^h	.05 (.90)	-1.38	1.48	-.19 (1.84)	-3.12	2.73
FRL Lunch ⁱ	.00 (1.00)	-.19	.19	.01 (.96)	-1.7	.18
ELL ^j	-.12 (.86)	-.41	.17	.15 (.78)	-.11	.42
Non-ELL ^k	.05 (1.05)	-.18	.29	-.07 (1.07)	-.31	.17
SPED ^l	-.63 (.62)**	-.99	-.27	-.57 (1.43)	-1.40	.25
Non-SPED ^m	.09 (1.00)**	-.11	.29	.08 (.89)	-.97	.26
Garfield ⁿ	.18 (1.03)*	-.08	.44	.35 (.85)***	.14	.57
Washington ^o	-.21 (.91)*	-.47	.04	-.42 (.99)***	-.70	-.14

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

Table 20

Means, Standard Deviations, and Confidence Intervals of Acuity C Outcomes in ELA and Mathematics by Demographic Variable

Group ^a	ELA			Mathematics		
	<i>M (SD)</i>	95% CI		<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>		<i>LL</i>	<i>UL</i>
Male ^b	.02 (1.01)	-.27	.31	.08 (.86)	-.17	.33
Female ^c	-.01 (.99)	-.26	.23	-.06 (1.08)	-.32	.21
Looped ^d	.26 (1.02)*	-.06	.58	.38 (.92)***	.10	.67
Did Not Loop ^e	-.15 (.95)*	-.37	.07	-.22 (.97)***	-.45	.01
White ^f	-.06 (1.26)	-.65	.53	.01 (1.05)	-.48	.50
Non-White ^g	.01 (.93)	-.18	.20	.00 (.98)	-.20	.20
Paid Lunch ^h	.09 (.94)	-1.40	1.58	-.38 (1.54)	-2.83	2.07
FRL Lunch ⁱ	.00 (1.00)	-.19	.19	.01 (.09)	-.17	.19
ELL ^j	-.08 (.85)	-.37	.21	.11 (.01)	-.20	.42
Non-ELL ^k	.04 (1.05)	-.20	.27	-.05 (1.03)	-.28	.18
SPED ^l	-.50 (.45)*	-.76	-.25	-.22 (1.27)	-.95	.51
Non-SPED ^m	.07 (1.03)*	-.13	.27	.03 (.95)	-.16	.23
Garfield ⁿ	.18 (1.06)*	-.09	.45	.31 (.95)***	.07	.55
Washington ^o	-.21 (.86)*	-.45	.03	-.37 (.91)***	-.63	-.12

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

Table 21

Means, Standard Deviations, and Confidence Intervals of December Categorical Survey Results by Demographic Variable. (SM = self-management. SE-ELA = self-efficacy in ELA. SE-M = self-efficacy in mathematics. PT = student perception of teacher. SA = social-awareness.)

Group	SM			SE-ELA			SE-M			PT			SA		
	95% CI			95% CI			95% CI			95% CI			95% CI		
	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL
Male ^b	3.53 (.80)**	3.29	3.76	3.34 (1.04)	3.03	3.64	3.54 (1.13)	3.21	3.87	3.91 (.77)	3.69	4.14	3.57 (.79)**	3.35	3.81
Female ^c	3.92 (.78)**	3.7	4.11	3.55 (.83)	3.35	3.75	3.74 (.77)	3.55	3.93	4.10 (.58)	3.96	4.24	3.94 (.61)**	3.79	4.09
Looped ^d	3.55 (.97)*	3.24	3.85	3.49 (.97)	3.19	3.80	3.70 (1.04)	3.37	4.02	4.04 (.70)	3.82	4.26	3.81 (.69)	3.60	4.03
Did Not Loop ^e	3.88 (.67)*	3.72	4.03	3.44 (.90)	3.22	3.65	3.64 (.89)	3.43	3.84	4.01 (.66)	3.85	4.16	3.77 (.73)	3.60	3.94
White ^f	3.77 (.91)	3.34	4.19	3.31 (1.14)	2.78	3.84	3.73 (1.08)	3.22	4.23	4.08 (.75)	3.73	4.43	3.79 (.69)	3.47	4.11
Non-White ^g	3.76 (.79)	3.59	3.92	3.49 (.88)	3.31	3.67	3.64 (.92)	3.46	3.83	4.01 (.66)	3.87	4.14	3.79 (.72)	3.63	3.93
Paid Lunch ^h	4.39 (.38)	3.79	4.99	4.19 (.66)	3.14	5.23	4.63 (.48)	3.86	5.29	4.50 (.28)	4.05	4.95	3.91 (.41)	3.25	4.56
FRL Lunch ⁱ	3.73 (.81)	3.58	3.89	3.43 (.93)	3.26	3.61	3.62 (.94)	3.45	3.80	4.00 (.68)	3.88	4.13	3.78 (.72)	3.65	3.92
ELL ^j	3.80 (.82)	3.52	4.07	3.43 (.81)	3.16	3.70	3.52 (.87)	3.23	3.81	3.92 (.80)	3.65	4.19	3.80 (.72)	3.56	4.05
Non-ELL ^k	3.74 (.81)	3.56	3.92	3.47 (.98)	3.25	3.70	3.72 (.97)	3.50	3.94	4.07 (.61)	3.93	4.20	3.78 (.71)	3.62	3.94
SPED ^l	3.24 (.93)**	2.70	3.77	2.95 (1.01)*	2.36	3.53	3.42 (1.04)	2.82	3.53	3.61 (.80)*	3.15	4.08	3.54 (.89)	3.02	4.05
Non-SPED ^m	3.83 (.77)**	3.68	3.98	3.53 (.90)*	3.35	3.71	3.69 (.93)	3.51	3.71	4.08 (.64)*	3.95	4.20	3.82 (.68)	3.69	3.96
Garfield ⁿ	3.73 (.93)	3.50	3.97	3.38 (1.01)	3.12	3.64	3.61 (1.05)	3.34	3.88	4.09 (.67)	3.92	4.26	3.79 (.69)	3.61	3.96
Washington ^o	3.78 (.63)	3.61	3.96	3.56 (.81)	3.33	3.78	3.72 (.80)	3.50	3.94	3.94 (.67)	3.75	4.13	3.79 (.74)	3.58	4.00

Note. CI = confidence interval; LL = lower limit, UL = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

Table 22

Means, Standard Deviations, and Confidence Intervals of February/March Categorical Survey Results by Demographic Variable. (SM = self-management. SE-ELA = self-efficacy in ELA. SE-M = self-efficacy in mathematics. PT = student perception of teacher. SA = social-awareness.)

Group	SM			SE-ELA			SE-M			PT			SA		
	95% CI			95% CI			95% CI			95% CI			95% CI		
	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL	M (SD)	LL	UL
Male ^b	3.87 (.70)	3.66	4.07	3.51 (.97)	3.22	3.79	3.81 (.95)	3.54	4.09	3.81 (.95)	3.88	4.27	3.69 (.74)*	3.48	3.90
Female ^c	4.08 (.74)	3.90	4.26	3.67 (.84)	3.46	3.87	3.87 (.93)	3.64	4.10	3.87 (.93)	3.99	4.34	3.97 (.68)*	3.80	4.14
Looped ^d	3.90 (.76)	3.66	4.14	3.70 (.79)	3.46	3.95	4.05 (.78)	3.81	4.30	4.02 (.69)	3.81	4.24	3.91 (.66)	3.70	4.11
Did Not Loop ^e	4.04 (.71)	3.88	4.20	3.54 (.95)	3.31	3.76	3.72 (1.00)	3.49	3.96	4.19 (.70)	4.02	4.35	3.82 (.75)	3.65	4.00
White ^f	3.98 (.80)	3.61	4.36	3.36 (1.00)	2.89	3.83	3.93 (1.26)	3.34	5.51	4.18 (.68)	3.87	4.50	3.74 (.66)	3.43	4.05
Non-White ^g	3.99 (.72)	3.84	4.14	3.65 (.87)	3.47	3.83	3.83 (.86)	3.65	4.01	4.12 (.71)	3.97	5.26	3.88 (.73)	3.73	4.03
Paid Lunch ^h	4.47 (.25)	4.08	4.86	4.13 (.75)	2.93	5.32	4.44 (.66)	3.29	5.48	4.54 (1.14)	4.31	4.77	4.5 (.43)	3.83	5.17
FRL Lunch ⁱ	3.97 (.73)	3.83	4.11	3.58 (.90)	3.41	3.74	3.83 (.94)	3.65	4.00	4.11 (.71)	3.98	4.25	3.83 (.72)	3.70	3.97
ELL ^j	4.01 (.77)	3.75	4.27	3.56 (.85)	3.28	3.85	3.92 (.88)	3.63	4.22	3.94 (.83)	3.66	4.22	4.84 (.63)	3.63	4.06
Non-ELL ^k	3.99 (.72)	3.82	4.14	3.62 (.92)	3.40	3.82	3.81 (.97)	3.60	4.02	4.21 (.62)	4.08	4.35	3.86 (.76)	3.69	4.03
SPED ^l	3.78 (.97)	3.21	4.34	3.05 (1.09)*	2.42	3.68	3.82 (1.13)	3.17	4.48	4.01 (.65)	3.63	4.39	3.83 (.79)	3.37	4.29
Non-SPED ^m	4.02 (.69)	3.88	4.16	3.68 (.84)*	3.51	3.84	3.85 (.91)	3.67	4.03	4.14 (.71)	4.00	4.28	3.86 (.71)	3.72	4.00
Garfield ⁿ	4.07 (.74)	3.88	4.26	3.50 (.98)	3.25	3.75	4.00 (.96)	3.75	4.24	4.12 (.70)	3.94	4.30	3.86 (.69)	3.68	4.03
Washington ^o	3.90 (.71)	3.70	4.09	3.72 (.77)	3.50	3.93	3.67 (.89)	3.42	3.92	4.14 (.70)	3.94	4.33	3.85 (.75)	3.64	4.06

Note. CI = confidence interval; LL = lower limit, UL = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

Table 23

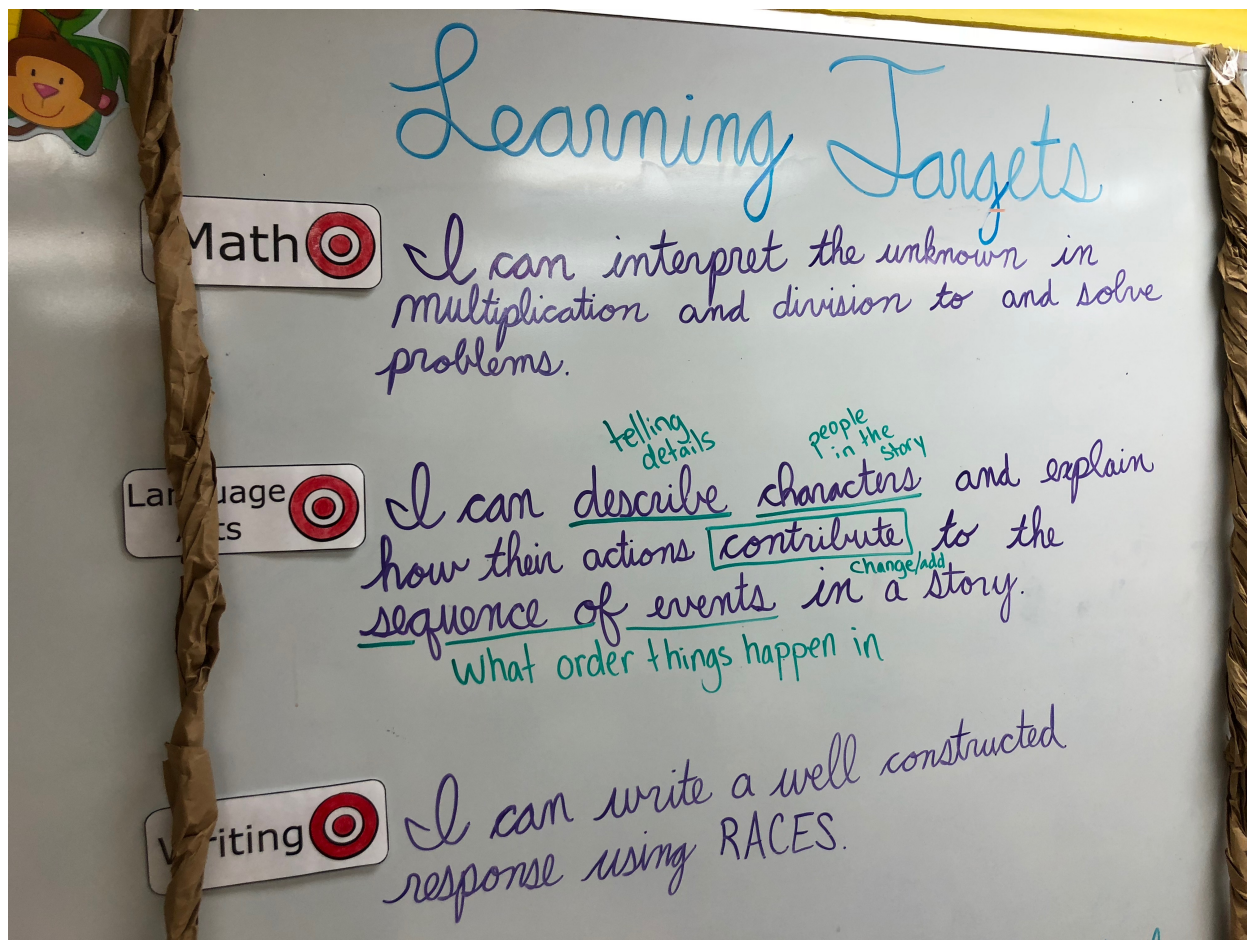
*Means, Standard Deviations, and Confidence Intervals of Teacher Perception of Student (TPS)
Survey Results by Demographic Variable*

Group ^a	TPS		
	<i>M (SD)</i>	95% CI	
		<i>LL</i>	<i>UL</i>
Male ^b	3.69 (1.14)*	3.35	4.02
Female ^c	4.10 (.96)*	3.86	4.33
Looped ^d	3.94 (1.16)	3.58	4.30
Did Not Loop ^e	3.92 (1.00)	3.68	4.15
White ^f	3.73 (1.34)	3.11	4.36
Non-White ^g	3.96 (.99)	3.76	4.17
Paid Lunch ^h	4.66 (.67)	3.61	5.73
FRL Lunch ⁱ	3.90 (1.06)	3.70	4.10
ELL ^j	4.11 (.91)	3.80	4.42
Non-ELL ^k	3.83 (1.11)	3.59	4.09
SPED ^l	3.76 (1.02)	2.18	4.35
Non-SPED ^m	3.95 (1.06)	3.74	4.16
Garfield ⁿ	3.73 (1.01)	3.45	4.01
Washington ^o	4.08 (1.07)	3.81	4.36

Note. CI = confidence interval; *LL* = lower limit, *UL* = upper limit. * $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed. ^a $N = 114$. ^b $n = 48$. ^c $n = 66$. ^d $n = 42$. ^e $n = 72$. ^f $n = 20$. ^g $n = 94$. ^h $n = 4$. ⁱ $n = 110$. ^j $n = 36$. ^k $n = 78$. ^l $n = 14$. ^m $n = 100$. ⁿ $n = 62$. ^o $n = 52$.

APPENDIX I

The following pictures were taken at Garfield and Washington Academies and illustrate how Learning Targets were posted in classrooms and on student work.



Some teachers hand wrote their targets on their board for the day in the style of an agenda.

Learning Targets

Reading

I can understand the "gist" of a text when reading it for the first time.

I can answer constructed response questions by referring to details from the text. →

Writing

I can **SHOW**, not tell in my writing.

I can add transition words and phrases in my narrative writing.

Math

Long term LT: I can demonstrate my understanding of division by solving multi-step real world problems.

I can use models to solve real world division problems with a remainder. Ex: $26 \div 3 = 8 \text{ R } 2$

I can understand and solve division problems with a remainder using array and area models. Ex: $26 \div 3 = 8 \text{ R } 2$

Science

I can demonstrate my understanding of place value and long division by solving division problems using a disc model.

$46 \div 2 = 23$

Tens	Ones
4	6
2	3

 Divisor 23, Dividend 46, long division problems.

CHOW

I can represent and solve division with up to a 4-digit dividend ex: $105 \div 4$

I can solve long division problems with a zero in the dividend or a zero in the quotient.

What MASTERY looks like

RACE

Some teachers expanded their targets to show the connectedness of supporting targets to a long term target with examples of mastery available for students to reference.

Division Unit:

I can represent and solve division problems with up to a 4 digit dividend.	1-22 LB	1-25 3	1-26		
I can demonstrate my understanding of division by solving multi-step real world problems.					
I can use RDW to show my mathematical thinking.	LT	1-9 2	2		
I can understand what the numbers in a problem represent.					
I can attend to precision (by showing my work correctly and clearly.)					

In which learning target do you feel like you have shown the most growth? Why?

Rate your grit throughout this unit (1-4) and explain your rating:

Other teachers organized their targets by unit and had students track their progress each day.

APPENDIX J

The following rubric is the CHOW (Character and Habits of Work) rubric used by Garfield and Washington Academies to define for students the behaviors of effective learners and school citizens. The seven targets are:

- I take responsibility for my working, learning and actions. This means I also own what I do and say, when I both do well and make mistakes.
- I am respectful to adults and my peers. This means my words, actions and tone of voice are polite, appropriate for school, and represent CHA well.
- I am independent. This means I can think, act and work on my own when appropriate.
- I have integrity. This means I am honest and do what is right even when it is hard.
- I am collaborative. This means I actively listen to others, contribute to ideas and conversations, and work with others to solve problems and create good work.
- I set goals, take action to meet them, and reflect on my progress.
- I have academic grit. This means I think through my academic problems, don't give up when things are hard, and make good things happen for myself and others.

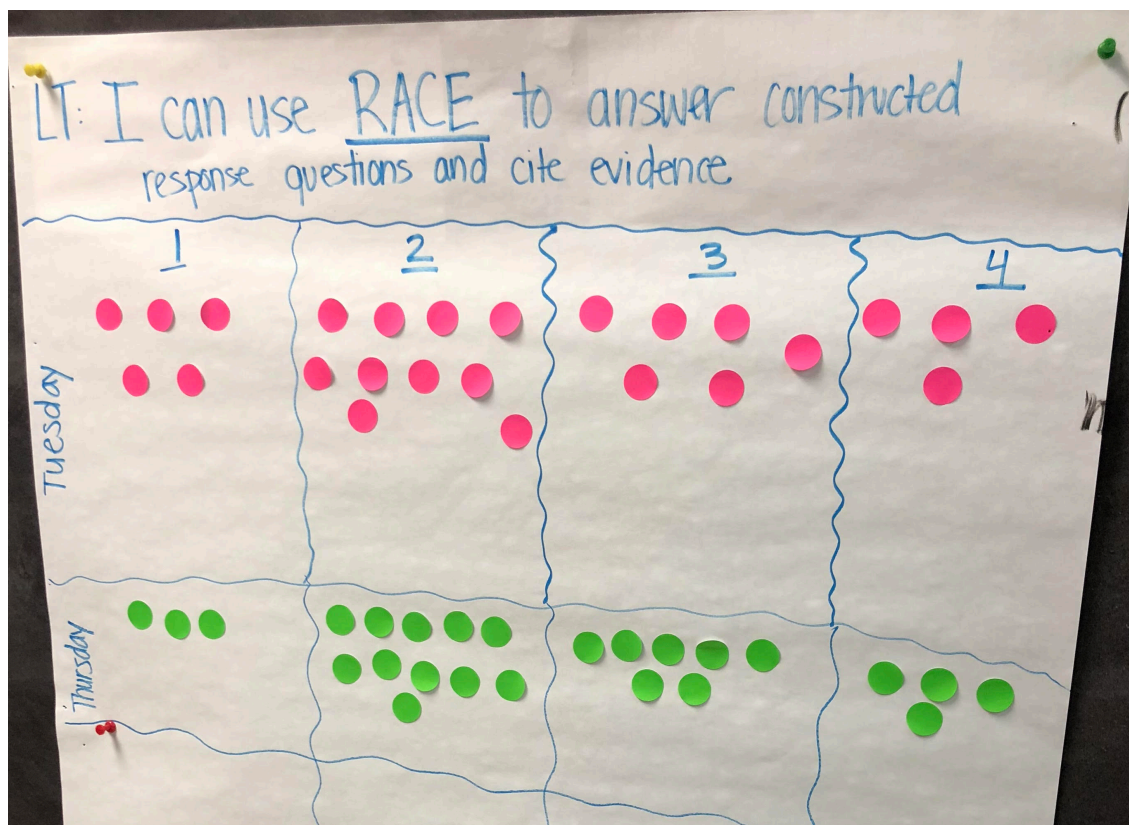
Beginning	Developing	Accomplished	Exemplary Leader
I take responsibility for my work, learning, and actions. This means I own what I do and say, when I both do well and make mistakes.			
<ul style="list-style-type: none"> Other people are usually to blame when I make mistakes in my work or get in trouble for my actions. My actions are reactions to others and not choices I make; as a result, I often blame others for the consequences of my actions. I struggle to understand how my actions affect my learning and success in class. My learning is the teachers' responsibility. I tend to just move on when I make mistakes or apologize if I'm made to. 	<ul style="list-style-type: none"> I hold myself responsible for my work and my actions at times or in certain situations. I recognize that some of my actions are choices I make; others are because people did or said things to me. I recognize that my actions affect my learning and success in class. My learning is mostly the teacher's responsibility. When I make a mistake I sometimes take steps to restore the trust and dignity of the people I'm close to and have <u>affected</u>. 	<ul style="list-style-type: none"> I hold myself responsible for my work and my actions. I recognize that my actions are choices I make, and no one else's. I recognize that my actions affect my learning and success in class and can articulate how. My learning is a shared responsibility between my teacher and me. When I make a mistake I take steps to restore the trust and dignity of the people I've <u>affected</u>. 	<ul style="list-style-type: none"> I consistently hold myself responsible for my work and actions, and can own them without prompting. I recognize that my actions are choices I make, and no one else's; I take full responsibility for them, sometimes publically, so that others can learn from me. I am actively in the "driver's seat" of my learning with the teachers as support, <u>and</u> an active participant in my own education, and help others do the same. I recognize that my actions affect my learning and success in class, can articulate how, and help others make connections between theirs. When I make a mistake I ensure that I take steps to restore the trust and dignity of the people I've <u>affected</u>, and help others realize they should do the same.
I am respectful to adults and my peers. This means my words, actions, and tone of voice are polite, appropriate for school, and represent CHA well.			
<ul style="list-style-type: none"> My words and behaviors don't affect others. It is hard to control my temper and tone. I often use disrespectful words and tone of voice in my communications with adults and/or peers. I struggle to use conventional, appropriate, and polite manners with other people. 	<ul style="list-style-type: none"> I am usually aware of how my words and behaviors affect others. I use disrespectful words and tone of voice in my communication with peers and adults infrequently but it still happens. I usually use conventional, appropriate and polite manners with other people. 	<ul style="list-style-type: none"> I am aware of how my words and behaviors affect others and try to be respectful. I regularly use respectful words and tone of voice in my communication with adults and peers. I regularly use appropriate and polite manners with other people in different types of situations. 	<ul style="list-style-type: none"> I am always aware of how my words and behaviors affect others and am quick to make adjustments to remain respectful. I use and encourage other students to use respectful words and tone of voice when communicating with adults and peers. I am always an exemplary representative of CHA's core values and use appropriate and polite manners with other people in all types of situations.
I am independent. This means I can think, act and work effectively on my own when appropriate.			
<ul style="list-style-type: none"> I struggle to manage my time well, and find myself off-task often. I need a lot of support to remain focused. I often rely on other's to do my work for me – I just don't get it or I don't want to do it. I struggle to make my own choices, and prefer instead to follow the actions of others, even if I know they aren't the right actions. 	<ul style="list-style-type: none"> I generally manage my time well, but I find myself distracted and off-task at times and need help to focus. I sometimes rely on other's to do my work for me. It's just easier that way at times. I struggle to make my own choices in certain situations, and prefer to follow the actions of others when I'm uncomfortable. I stop though when I know it's wrong. 	<ul style="list-style-type: none"> I manage my time well, and am on-task when I have work to complete. I regularly work independent of others; it is important that I learn information and skills for myself (not necessarily by myself). I make my own choices based on what I think in almost all situations and can regularly solve problems by myself. 	<ul style="list-style-type: none"> I manage my time well and always pride myself on doing my own work. I make sure I stay productive by working on other things when I'm finished with my work. I am a model of independence, and work well individually so that I can master the information and skills for myself, but I also know how to work with and help others. I am an independent thinker, encourage others to be independent when necessary, and help them reach independence by being a leader inside and out of school.

I have integrity. This means I am honest and do what is right even when it is hard.			
<ul style="list-style-type: none"> I have a tendency to lie when it is the easiest choice to make, or when it is an advantage for me. I struggle to make good decisions when faced with difficult situations, especially when the right thing is the harder choice to make. I put myself and my own interests first even if it means what I'm doing is wrong or selfish. 	<ul style="list-style-type: none"> I sometimes lie or exaggerate about small, inconsequential things. It is really hard for me but I usually do the right thing in difficult situations. I sometimes make bad choices when the right ones are really hard. I sometimes put others before my own self-interest but I still put myself before what I know to be right at times. 	<ul style="list-style-type: none"> I choose to tell the truth, even when it is hard. I always try to do the right thing even in difficult situations when it is hard. I put others in front of my own self-interests at appropriate times, realizing I need to take care of me too. 	<ul style="list-style-type: none"> I tell the truth, and promote a truthful environment for myself and my peers. I always try to do the right thing, and encourage others to do the right thing especially when it is hard. I put others before myself at the appropriate times, making sure we are both better for my choice. In doing so, I help foster a selfless and supportive environment while making sure I'm taken care of as well.
I am collaborative. This means I actively listen to others, contribute to ideas and conversations, and work with others to solve problems and create good work.			
<ul style="list-style-type: none"> If I don't get along with someone in a group I don't make the effort to listen to or work with them. I am pretty quiet or do my own thing and don't really offer suggestions and ideas or help much when working with others. I don't really have an interest in solving problems with others in a group and I am not concerned with contributing to the work of the group. My behaviors don't really matter to or affect others in the group. 	<ul style="list-style-type: none"> I sometimes try to get along with people in my group. I get distracted and struggle to actively listen while others contribute ideas and suggestions during group conversation. I sometimes work with others to solve problems and try to help get the work done. I sometimes respect others opinions and show interest in considering their ideas. 	<ul style="list-style-type: none"> I always try to get along with the people in my group. I try to actively listen to others in the group by having appropriate body language (eye contact, head gestures). During conversation, I offer advice and points to help further the group's goal. I share my thoughts with the group, offer suggestions on how to complete the group's task, and help get the work done. I respect other people's opinions in my group, even if I do not share them. 	<ul style="list-style-type: none"> I actively work to get along with all of the people in my group. I listen to others in the group by having appropriate body language (eye contact, head gestures) and politely encourage others to stay on task and do the same. I encourage others to speak their minds, and offer advice and points to help further the group's goal. I share my thoughts with the group, and offer suggestions on how to complete the group's task. I also build on other's ideas to make our ideas stronger, and make sure we get our work done with quality. I help create a respectful environment for my group in which everyone's opinions are respected.
I set goals, take action to meet them, and reflect on my progress.			
<ul style="list-style-type: none"> I don't think it's important to set goals so I don't really set them unless I'm made to. If I have goals, I don't follow through with them or take appropriate actions that allow me to accomplish them. The choices I make don't really affect my ability to reach goals. I don't really reflect on my goals or make revisions to my action plans. 	<ul style="list-style-type: none"> I sometimes set goals but I usually set just short-term goals for myself, or I set a long term goal with no supporting goals to help me get there. I sometimes take advantage of situations that allow me to accomplish my goals but I don't go out of my way to complete them. I'm sometimes intentional about making choices that help me reach my goals. I only reflect on my goals or make revisions to my action plans when I'm told or reminded to. 	<ul style="list-style-type: none"> I regularly set both long and short term goals, and work hard to accomplish them. I regularly take advantage of situations that allow me to accomplish my goals and make opportunities happen for myself. I choose courses of action that have a positive impact on my ability to accomplish my goals. I regularly reflect on my goals, and revise my action plan to help me meet them. 	<ul style="list-style-type: none"> I always set meaningful long term goals and intentional short-term goals, and work hard to accomplish them. I motivate my peers and help them to set goals for themselves. I always take advantage of situations that allow me to reach my goals, and make opportunities happen for myself and others. The courses of action that I chose for myself consistently have a positive effect on my ability to accomplish my goals. I encourage others to make good choices and help advise them on a course of action as well. I regularly reflect on my goals and accomplishments, and use my choices and actions to set future goals that build on what I have already achieved. I motivate my peers and help them to reflect on their goals and action plans.
I have grit. This means I think through problems, don't give up when things are hard, and make good things happen for myself and others.*			
<ul style="list-style-type: none"> I usually give up when things become difficult to accomplish. I usually just react rather than think through and use my problem solving skills to help me tackle tough situations. I just let things happen around me and rarely make opportunities for myself or others. 	<ul style="list-style-type: none"> I am able to push forward when things are hard but tend to give up when they get really difficult or uncomfortable, especially if I don't see their importance. I sometimes use my problem solving skills to help me tackle tough situations but I also let others solve my problems for me, or just take the consequences for not solving them. I sometimes make good things happen for myself. 	<ul style="list-style-type: none"> I regularly push forward when things are difficult. Even though I'd like to give up, I persevere to get through. I am determined to solve problems, even in tough situations. I do this in a positive and productive way, and reach out to others for support when I need it. I regularly try my best to make good things happen for myself and others. 	<ul style="list-style-type: none"> I regularly push forward when things are difficult. My grit and encouragement motivates those around me to also persevere. I am determined to solve problems especially in tough situations, and I do it in a proactive, positive, and productive way. I reach out to others for support when I need it and I am a source of support for others to persevere. I always try to make good things happen for myself and others, and my positive outlook encourages my peers to do so as well.

*Grit should be distinguished between home-life and academic perseverance.

APPENDIX K

The following pictures were taken at Garfield and Washington Academies and illustrate how teachers used generic public data charts and graphs to compare students' level of development towards various classroom objectives. This evidence relates to the *using peer models for self-evaluation or comparison* CASE indicator under the Vicarious Experiences subscale.



Most of the 10 classrooms had charts such as this on their walls. Students were asked to place a sticker on their level of development towards the task or objective. Typically the scale represented their level of understanding or mastery where a 1 represented Beginning, 2 was Developing, 3 was Accomplished and 4 was Exemplary.

CHOW TARGET PROGRESS TRACKING

LT: I can track my own CHOW progress.

As a result, CHOW goals will become more prominent in all areas of my learning growth.

CHOW	1	2	3	4	5	6	7	8	9	10
I take responsibility for my work, learning, and actions. This means I take ownership of what I do and say, when I both do well and make mistakes.	😊	😊	😊							
I am respectful to adults and my peers. This means my words, actions, and tone of voice are polite, appropriate for school, and represent CHS well.	😊	😊	😊	😊	😊					
I am independent. This means I can think, act and work on my own when appropriate.	😊									
I have integrity. This means I am honest and do what is right even when it is hard.	😊	😊	😊							
I am collaborative. This means I actively listen to others, contribute to ideas and conversations, and work with others to solve problems and create good work.	😊	😊	😊	😊	😊	😊	😊	😊	😊	😊
I set goals, take action to meet them, and reflect on my progress.										
I have grit. I think through problems and don't give up when things are hard.										

10 stickers in any CHOW Target= bean bag or teacher desk for the day

NAME: _____ # 9

5th Grade Class

Great Work!

CHOW TARGET PROGRESS TRACKING

LT: I can track my own CHOW progress.

As a result, CHOW goals will become more prominent in all areas of my learning growth.

CHOW	1	2	3	4	5	6	7	8	9	10
I take responsibility for my work, learning, and actions. This means I take ownership of what I do and say, when I both do well and make mistakes.	😊	😊	😊	😊						
I am respectful to adults and my peers. This means my words, actions, and tone of voice are polite, appropriate for school, and represent CHS well.	😊	😊	😊	😊	😊					
I am independent. This means I can think, act and work on my own when appropriate.	😊									
I have integrity. This means I am honest and do what is right even when it is hard.	😊	😊	😊							
I am collaborative. This means I actively listen to others, contribute to ideas and conversations, and work with others to solve problems and create good work.	😊	😊	😊	😊	😊	😊	😊	😊	😊	😊
I set goals, take action to meet them, and reflect on my progress.										
I have grit. I think through problems and don't give up when things are hard.										

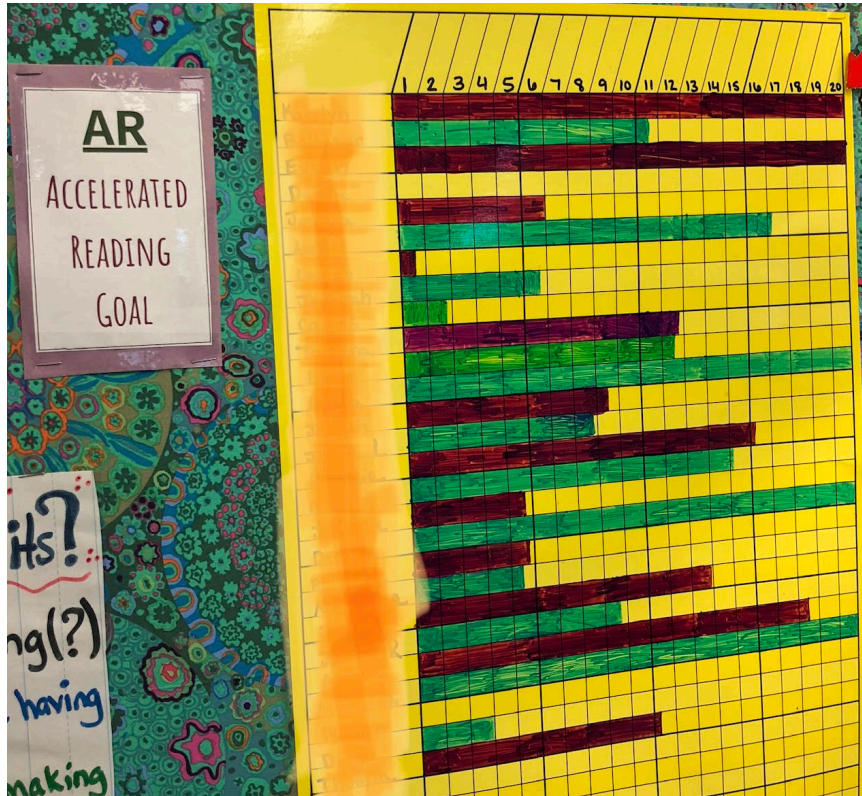
10 stickers in any CHOW Target= bean bag or teacher desk for the day

NAME: _____ # 10

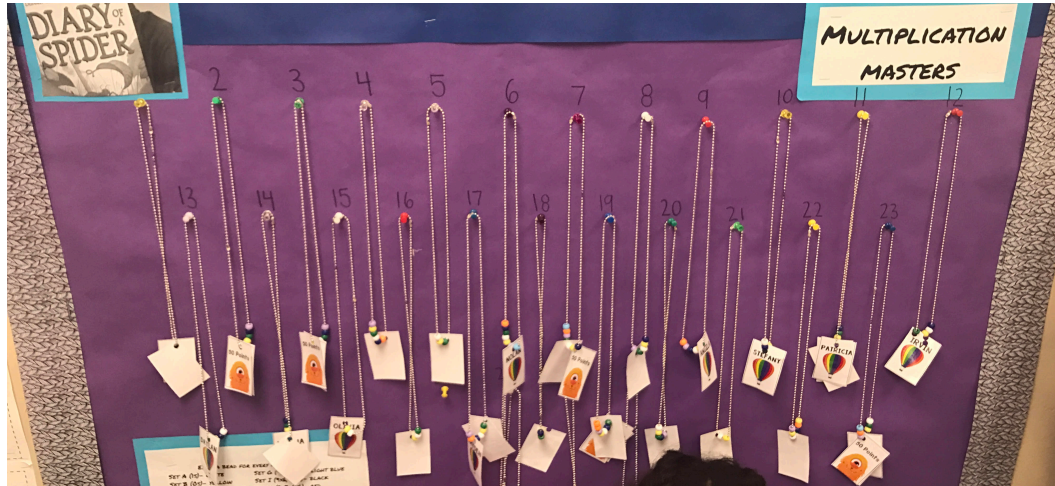
5th Grade Class

Great Job!

Students tracked their progress towards their Character and Habits of Work (CHOW) goals. The implementation varied from classroom to classroom. Students in some classrooms indicated they self-tracked and put stickers on their own charts, students in others indicated they only received stickers from the teacher. Only one classroom had a student respond that they also gave stickers to each other if they recognized a peer exemplifying the behaviors listed on the chart.



Students in this classroom tracked their progress towards meeting their Accelerated Reader reading goals for the semester.



One classroom had necklaces where students received badges for successfully passing their math facts timed tests. A student I conversed with said they pushed her to try harder when she saw her friends get more badges and pass more tests than her. She also reflected, “I like getting my badge. I like counting them with my friends during [indoor] recess... my friend doesn’t have as many as me and I like that I can help [teach] her.”